Noise and epigenetic inheritance of single-cell division times influence population fitness

Authors

Bram Cerulus[1,2,a], Aaron M. New[1,2,3,a] , Ksenia Pougach[1,2] and Kevin J. Verstrepen[1,2,*]

Affiliations

1 VIB Laboratory of Systems Biology, Gaston Geenslaan 1, B-3001 Leuven, Belgium.

2 CMPG Laboratory of Genetics and Genomics, Department M2S, KU Leuven, Gaston Geenslaan 1, B-3001 Leuven, Belgium.

3 EMBL-CRG Systems Biology Unit, CRG and UPF, Barcelona 08003, Spain

Contact Information

*Correspondence: Kevin.Verstrepen@biw.vib-kuleuven.be

Additional Title Page Footnotes

a These authors contributed equally to this work. Co-first authors.

Summary

The fitness effect of biological noise remains unclear. For example, even within clonal microbial populations, individual cells grow at different speeds. Although it is known that the individuals' mean growth speed can affect population-level fitness, it is unclear how or whether growth speed heterogeneity itself is subject to natural selection. Here, we show that noisy single-cell division times can significantly affect population-level growth rate. Using time-lapse microscopy to measure the division times of thousands of individual *S. cerevisiae* cells across different genetic and environmental

backgrounds, we find that the length of individual cells' division times can vary substantially between clonal individuals, and that sublineages often show epigenetic inheritance of division times. By combining these experimental measurements with mathematical modeling we find that, for a given mean division time, increasing heterogeneity and epigenetic inheritance of division times increase the population growth rate. Furthermore, we demonstrate that the heterogeneity and epigenetic inheritance of single-cell division times can be linked with variation in the expression of catabolic genes. Taken together, our results reveal how a change in noisy single-cell behaviors can directly influence fitness through dynamics that operate independently of effects caused by changes to the mean. These results not only allow a better understanding of microbial fitness, but also help to more accurately predict fitness in other clonal populations such as tumors.

Introduction

The fitness of a population depends on the reproductive performance of each of the individuals in the population. However, populations can be very heterogeneous, for example due to differences in so-called life-history traits such as the kinetics of reproduction, age and mortality [1]. Interestingly, many studies have revealed that even clonal populations in homogenous environments can show substantial levels of heterogeneity [2–12]. For example, microbial populations that are growing in the exponential phase show a considerable heterogeneity in the length of doubling (division) times (DTs) between different cells [13–19]. Similarly, cell-to-cell growth rate heterogeneity also occurs in other clonal cell populations such as tumors [20–22]. While such biological noise in growth behavior has often been interpreted to have a direct negative impact on population-level growth rate [2, 23, 24], this effect has rarely been quantified and analyzed in detail [25].

Gene expression noise, *i.e.* stochastic variability in gene expression, is a key factor believed to contribute to differences between cells in a clonal population [2, 26–29]. Gene expression noise can be a disadvantageous imperfection, for example when a robust and precise physiological response to an environmental change is required to maintain high fitness [30–33]. However, phenotypic heterogeneity arising from gene expression noise can also be part of a bet-hedging strategy, for example by creating subpopulations that are prepared for changing or adverse conditions, often at a fitness cost in the current environment [7, 9, 34–37]. Similar to noise in single-cell growth, the direct quantitative impact of gene expression noise on population-level fitness has received little attention [7, 9, 34, 38–40].

Here, we use a combination of modeling and experimentation to investigate in detail how key life-history traits including mean, variation and epigenetic inheritance of DTs affect a population's fitness. Using time-lapse microscopy, we acquired single-cell DT distributions for a diverse range of genetically distinct *Saccharomyces cerevisiae* yeast strains growing exponentially on medium supplied with different carbon sources. Measurements of mother and daughter cell DTs indicate that

certain strain/medium combinations yield noisy DT distributions with substantial epigenetic inheritance of DT within sublineages. A stochastic model reveals that DT variability within mother and daughter fractions and epigenetic inheritance of DT increase the population growth rate considerably beyond the predictions of a simple deterministic model. We show that this surprising result can be explained by a complex evolution towards a steady-state distribution of single-cell growth rates within the population. Finally, using a reverse-genetics approach, we show how changes in gene expression of catabolic genes can contribute to noisy single-cell growth behaviors. Together, our results show how variability in life-history traits across clonal individuals can sometimes counter-intuitively affect population-level growth rates.

Results

Measuring single-cell DTs using automated live-cell microscopy

To investigate growth at the single-cell level, we used automated time-lapse microscopy to measure key life-history traits including mean, variance and epigenetic correlations of DTs of single-cells in clonal populations of exponentially growing yeast cells (Experimental Procedures). We acquired time-lapse growth records for 10 genetically distinct yeast strains growing in up to 7 different growth media. By varying carbon source, we were able to analyze single-cell DTs across a wide range of population growth rates (N=41 experiments; see Dataset S1). These population growth rates were measured on solid media by tracking microcolony growth, and in liquid media using a colony-counting assay (Supplemental Experimental Procedures and Dataset S3).

For each experiment, we analyzed the time-lapse movies by tracking growth of individual cells within 16-113 microcolonies, yielding a total number of more than 5500 single-cell DTs (Experimental Procedures and Dataset S1). Figure 1A and Movie S1 show how we measured the single-cell DTs. In our analyses, we distinguish between mothers (cells that have already completed a bud), and daughters (newborn cells that have not completed a bud yet). For each experiment, this analysis yields mother and daughter DT distributions, which are represented in Figure 1B.

DT variability and epigenetic inheritance are condition- and strain-dependent

Several trends emerged from this dataset. First, it is clear that although there is considerable variability in mean single-cell DTs across strains and environments, mean daughter and mother DTs display a striking linear relationship (Figure 2A; R^2 = 0.881, p<2.2e-16). We also find that the large increase in mean daughter DTs at slower growth rates is mostly accounted for by an increase in the length of the unbudded period of the daughter cells (Figure 2B).

In order to statistically summarize the variability (or noise) in DTs for mother and daughter cells, we calculated the coefficient of variation (CoV; standard deviation (SD) / mean). This trait is partially correlated with the mean DT (Figure 2C). However, respectively 37% and 75% of the variability in mother and daughter CoV is left unexplained by the mean DT, which implies that DT noise in strains varies considerably independently of the mean. For example, the strains Y55 and BC187 growing in galactose have very different levels of DT CoV at similar mean DTs (BC187 vs. Y55; mother DT CoV = 0.201 vs. 0.098 and daughter DT CoV = 0.215 vs. 0.139; Figures 1B and 2C).

Apart from raw DT values, our dataset also allowed us to follow how DTs were correlated between mothers and daughters (Figure 1A and Figure S1A). Using these genealogical relationships, we investigated if there is epigenetic inheritance of the DT length by measuring correlations between individual DTs within lineages ([41, 42] and Figure S1). We find that on average the strongest DT correlation exists between a mother and her most recently born daughter cell (Figure 2D; R^2=0.2517). In contrast, the average correlation between consecutive DTs of a given mother was much weaker (Figure S1D; R^2=0.0153). These traits are independent of mean and CoV in DT (Dataset S4). Notably, across our 41 experiments, certain strain/condition combinations displayed considerably higher DT correlations than others (Dataset S1). For example, for BY/S288c growing in palatinose we find that the R^2 of both genealogical relationships is higher than 0.4 (see below).

Using this dataset, we used linear regression to investigate which DT characteristics best explain population-level growth. We found that a simple linear model based only on the DT means is highly

predictive of population-level growth rate (Dataset S4; R^2 = 0.880, p<2.2e-16). The accuracy of this model is not significantly improved by adding the DT CoV and DT correlations as parameters (Dataset S4).

An individual-based model that combines single-cell variance and epigenetic behavior to predict population-level growth rates.

To gain insight into how the mean, noise, and epigenetic inheritance of single-cell DTs affect population-level growth behavior, we used mathematical analysis and simulation. Previously, an elegant single-cell model for budding yeast population growth was proposed by Hartwell & Unger [43]. In this deterministic model, mothers and daughters are modeled to grow at a fixed DT (measured as the arithmetic mean of the empirical DT distributions).

Deterministic Model:

$$mother\ DT = \mu_{\text{mother}}$$

$$daughter\ DT = \mu_{\text{daughter}}$$

This model can be solved analytically, allowing the prediction of population growth rate during steady-state growth ([43, 44] and Experimental Procedures).

Importantly, this deterministic model considers only the arithmetic mean of mother and daughter subpopulations and does not take into account the inter-individual variability. To explore the effect that single-cell variability in DTs could have at the population-level, we developed an individual-based model of population growth that accounts for variance and epigenetic inheritance of cellular DTs. In our model, all cells divide at variable DT lengths, which are randomly assigned to them by sampling from a distribution specific for the mother and daughter cells (Figure 3A). For mathematical simplicity, we describe the model in terms of the parameters of a normal distribution that were fitted to empirical DT measurements, which showed a good fit for the data (Dataset S2 and Supplemental

Experimental Procedures). However this analysis can be extended to other distributions, including simply the empirical DT distribution (Figure S4 and Supplemental Experimental Procedures).

Stochastic Model (Without epigenetics):

$$mother\ DT = N(\mu_{\text{mother}}, {\sigma_{\text{mother}}}^2)$$

$$daughter\ DT = N\left(\mu_{\text{daughter}}, {\sigma_{\text{daughter}}}^2\right)$$

Such a stochastic model does not yet take into account the epigenetic inheritance of DTs, since DTs are distributed randomly and independently of previously assigned DTs. To include the effect of epigenetic inheritance of DTs on the population growth rate, we simulate that the choice of new DTs also depends on previous mother DTs (Figure 3A and Figure S2). For each cellular lineage in the simulation, previously assigned DTs are used as an input to determine new DTs (the output). The set of equations used to determine new DTs comprise a transfer function that returns a series of output DTs matching the empirical DT distribution. Further, the transfer function's parameters can vary the strength of the correlation (R^2) between the input and output distributions.

Stochastic Model (With epigenetics):

$$mother\ DT = DT_{output,zscore} * \sigma_{mother} + \mu_{mother}$$

$$daughter\ DT = DT_{output,zscore} * \sigma_{daughter} + \mu_{daughter}$$

$$DT_{output,zscore} \sim \mathcal{N}\left(A \times DT_{input,zscore}, \sigma^2_{transfer}\right) \quad with\ \sigma_{transfer}, A \in [0,1]$$

These equations introduce two new parameters that reflect the extent of DT correlation: $\sigma^2_{transfer}$ and A. In the Supplemental Information, we show that for certain parameter combinations, when provided with a series of normally distributed input DTs, these functions return a series of normally distributed output DTs (Figure S2B). Each combination of A and $\sigma^2_{transfer}$ leads to a certain correlation (R^2) between the input and output distributions. In this way, the R^2-values of the

experimentally determined DT correlations are used to choose the parameters A and $\sigma^2_{transfer}$ for strain/condition specific modeling.

To obtain population growth rate estimates for our stochastic model, we ran Monte Carlo simulations of population growth based on our empirically measured DTs (Supplemental Experimental Procedures). We ran the model under different assumptions: 1) strictly deterministic, 2) strictly stochastic (without epigenetic effects), and 3) including both DT variability and epigenetic behavior (Figure 3A). We find that all models accurately predict experimentally-measured population growth rates made on solid medium and in liquid culture (Figure 3B,C and Figure S3). In other words, the deterministic model predicts experimentally-measured population-level growth with a very similar accuracy as the stochastic and epigenetic models, suggesting that the DT mean is the strongest determinant of population growth rate across our dataset of growth measurements in a wide range of environments. Consistent with this, as previously mentioned, we found that across all experiments in our dataset, the mean DT alone statistically explains 88% of population-level growth rate (Dataset S4). However, the results in the next paragraphs indicate that for certain environments and strains, inter-individual variation and epigenetic inheritance of DTs become important and also significantly affect the population-level growth.

Heterogeneous single-cell growth affects population-level fitness

Our growth simulation model allowed us to address two key questions about the potential population-level effects of our single-cell observations: 1) What role does variance in single-cell DTs play in population fitness? and 2) what role does epigenetic inheritance in single-cell DTs play? Contrasting our model's predictions of population-level growth rate with the prediction of the classic deterministic model allowed us to measure the effect that these factors can have on fitness.

To attack the first question of how variability in DTs can affect population-level behavior, we compared our stochastic model with the deterministic model across our 41 experiments. This analysis reveals that a deterministic model can underestimate population growth rate by up to 4-7%,

depending on the type of distribution used in the simulation (see Figure 4A and S4A). Further analysis indicates that DT CoV explains nearly all of this growth rate increase (R^2 = 0.96, p<2.2e-16). Even when controlling for the weak covariation between DT mean and DT CoV (Figure 2C), variation in single-cell DT alone can explain 63% of this growth rate increase (Dataset S4, Supplemental Experimental Procedures, p<2.2e-16).

We used stochastic simulations to systematically investigate how these aspects of DT variation affect population-level growth. Consistent with our statistical observations, these analyses indicated that the DT CoV has a large effect on the population growth rate across the wide range of values in our experimental measurements (Figure 4B). By contrast, although we find that skewness and kurtosis can have large population-level growth rate effects, such effects are negligible in the parameter space that we observe in our dataset (Figure S4A,B).

The results above indicate that, given a constant mean DT, stochastic variation in DTs increases growth rate. To understand how this counter-intuitive effect occurs, we first have to consider the growth rate of each single cell, which is given by ln(2)/DT. Importantly, given a DT distribution with mean = μ and SD > 0, the mean of the corresponding growth rate values will actually be higher than the growth rate of the mean DT = ln(2)/μ. This effect is more generally known as the Jensen's inequality, which states that the mean of a set of values that have undergone a convex transformation f(x), in this case growth rate=ln(2)/DT, is equal to or higher than the same transformation of the mean of these values (or mean[f(x)] ≥ f(mean[x])). Therefore, for a given mean, adding noise to a single-cell DT distribution has the potential to increase the population growth rate, since on average cells are assigned faster growth rates over the course of each doubling.

However, Jensen's inequality alone is not sufficient to explain how population-level growth rate changes with increasing DT noise. We also have to take into account that the single-cell DT distribution evolves with time to reach a steady-state that is enriched with slowly growing cells (Figure S4D). This is because individuals with short DTs finish their doubling more quickly than cells

with a long DT, while new individuals have equal chances of growing with a long or a short DT. In the Supplemental Information, we show that it is the arithmetic average of the single-cell growth rates at steady-state that equals the population-level growth rate. Therefore, the enrichment of slower-growing individuals during steady-state has the potential to reduce, or even counteract, the population growth rate increase which is predicted by Jensen's inequality. Combining both effects, our stochastic model shows that there is a net population growth rate increase, which is however lower than the arithmetic mean of the sampled growth rate (= ln[2]/DT) distribution (Figure S4C).

The effect of increased DT noise on population-level growth can be non-trivial, especially given that the variance and mean of single-cell DTs can be largely independent phenotypic traits (Figure 2C). The results imply that a change in DT CoV can yield a considerable difference in fitness and thus a significant selective advantage over evolutionary timescales. Indeed, a simple simulation using the stochastic model shows that, when the mean DT is kept constant, mutants with more noisy DT distributions have a selective advantage compared to a population with less noisy DT distributions. Moreover, these mutants reach fixation at similar frequencies as mutants that have the same fitness advantage by having a shorter DT mean, but the same DT noise (Supplementary Experimental Procedures). Together these results are at odds with the common intuition that growth noise is always detrimental for the population-level fitness.

Epigenetic inheritance of single-cell DTs further increases population-level fitness

To investigate the effect of epigenetic DT inheritance on population growth rate, we compared the predicted population growth rates from the stochastic-epigenetic model with those of the strictly stochastic model (Figure 5). For each specific strain/condition combination, we modeled two experimentally observed DT correlations, including the correlation between consecutive mother DTs and the correlation between the DTs of a mother and her most recently born daughter (Figure 5A and Figure S2C).

Viewed across our dataset, epigenetic inheritance has only a small effect on population growth rate compared to predictions made when there is no individual DT correlation (Figure 5B). However, some strain/condition combinations show a significant population-level growth rate increase when epigenetic effects are considered, especially for strain/condition combinations that show both high single-cell DT variability and high epigenetic DT inheritance (Figure S5). We illustrate this in Figure 5C with a fitness landscape of the strain BY/S288c growing in palatinose. The growth rate-increasing effect of epigenetic DT inheritance can be intuitively understood by its effect on the steady-state DT distribution. Even though cells with a short DT finish their doubling more quickly, because of the epigenetic factor, they are more likely to give rise to fast growing cells themselves. This diminishes enrichment for slowly-diving cells, leading to a steady-state DT distribution containing more quickly-dividing cells than under the assumption of a purely stochastic model (Figure S4D,E).

Epigenetics can amplify the heterogeneity of growth rates between small populations

We found that single-cell DT variability and epigenetic DT inheritance can affect population-level growth rates at large population sizes (Figures 4 and 5). Likewise, we expected that these parameters could also considerably affect the variability of growth rates between different (isogenic) populations, especially at small population sizes when the variability in single-cell DTs might not yet be averaged out [7, 9, 24, 45].

To test for these effects, we used a non-parametric statistical comparison between our model's predictions and empirical measurements made at small population sizes (Supplemental Information). These analyses revealed that for strain/condition combinations where the mother-daughter DT correlation is low ($R^2 < 0.35$), epigenetics plays a small role in microcolony variability. However, for strain/condition combinations where the mother-daughter DT correlation is high ($R^2 > 0.35$), simulations of microcolony growth rate distributions that account for epigenetic effects are consistently closer to the measured values (Figure S6B; Supplemental Information). Together these results indicate that in many cases, microcolony growth variability is simply the result of stochastic

variability in single-cell DTs. However, in other cases where epigenetic effects are strong, this can lead to further increases in microcolony growth rate variability.

Overexpressing the genes required for growth reduces heterogeneity and epigenetic inheritance of single-cell DTs

We set out to identify a molecular mechanism that might underlie stochastic and epigenetic single-cell DT variability. To explore this, we examined gene expression and single-cell growth in palatinose, a condition that gave rise to both a high degree of single-cell DT variability and high epigenetic DT inheritance (Figure 5A). We reasoned that expression of the genes necessary for growth on this sugar might affect the observed growth variability. To grow fermentatively on palatinose, *S. cerevisiae* needs to transport the sugar through the alpha-glucoside transporter Mal11p and cleave the sugar intracellularly using an alpha-1,6-glucosidase enzyme (enzyme encoded by two paralogous genes, *IMA1* and *IMA5*) [46–48]. The expression of *MAL11* and *IMA1* is induced by two transcriptional activators (*MAL13* and *YFL052W*), both of which are essential for growth on palatinose [46]. The regulatory gene network is visualized in Figure 6A.

We constructed a strain bearing fluorescently tagged *IMA1(*-yECitrine) and *MAL11(*-mCherry) alleles and measured expression of these two genes with flow cytometry. The results indicate that while both genes are on average highly expressed in palatinose, they exhibit high expression noise (Figure 6B). Moreover, expression levels of *MAL11* and *IMA1* show a significant correlation at the single-cell level (R^2=0.37±0.01 of log-transformed fluorescence; $p<0.001$). We hypothesized that noise in gene expression of Ima1p could be reduced by overexpressing *MAL11*, since this transporter is part of a positive feedback loop in the regulatory network (Figure 6A). Indeed, overexpression of *MAL11* in the fluorescently tagged strain (using an extra untagged copy of *MAL11* expressed from a constitutive *GPD* promoter) results in reduced network noise, with much more homogeneous Ima1p expression levels (Figure 6B), leading to an approximately 7-fold reduction in single-cell expression noise of Ima1p-mCitrine (gene expression noise $\frac{\sigma^2}{\mu^2}$ [26] lowers from 0.549 to 0.075).

Next, we investigated if these changes in the expression of the palatinose utilization pathway would affect DT variability and epigenetic inheritance by measuring single-cell growth using time-lapse microscopy. The results show that both single-cell DT variability and epigenetic DT inheritance are strongly reduced by overexpression of *MAL11* compared to WT control cultures (Figure 6C; p= 0.011 and 0.052 for mother and daughter DT CoV, respectively). These changes are comparatively large as well: across all measurements made in this study, variability in daughter DT distributions (the residuals of the plot of CoV vs. mean; Figure 2C) shifts from the 84rd percentile to the 12th percentile (p=0.052) and variability in mother DT distributions is reduced from the 88th percentile the 60th percentile (p=0.011 and Supplemental Information and Dataset S4). Finally, we find a strong reduction in epigenetic DT inheritance; the correlation (R^2) between consecutive mother DTs is reduced from R^2 = 0.456 to 0.114 (p=0.010), and the correlation between mother and daughter DTs is lowered from 0.452 to 0.258 (p=0.045). In contrast, the effect of *MAL11* overexpression on the mean mother DTs and mean daughter DTs was weak and opposite for mothers and daughters (mean mother DT reduced from 1.86h to 1.71h; mean daughter DT increased from 3.35h to 4.07h). Taken together, these results indicate that a single genetic modification can affect the variance and epigenetic features of single-cell DTs with inconsistent effects on the mean. Further, they show that the expression of genes required for growth can modulate DT variability and epigenetics.

Discussion

Together, our observations indicate that three life-history traits of cellular growth – mean, variance and epigenetic inheritance of DT – are naturally variable, largely independent, and genetically influenced traits. Furthermore, our mathematical models and simulations indicate that these traits have the potential to significantly affect population-level fitness. Most intuitively, population-level growth rate is affected by the mean DT. However, for a given mean DT, individual-level variation in DT can further influence population-level growth, with higher variance counter-intuitively increasing fitness (Figure 4). Finally, for a given level of mean and variation, increasing epigenetic inheritance in

DTs can further increase population-level growth (Figure 5). We show that the effect of DT variation and epigenetic inheritance occurs through a complex evolution towards a steady-state distribution of single-cell growth rates within the population. These results are in line with results obtained by Tănase-Nicola & ten Wolde (2008), who used mathematical models to show that Gaussian noise in the instantaneous single-cell growth rate could increase population growth rates.

Given that increased DT variation and epigenetic inheritance of DTs have the potential to increase population-level growth rate, one might wonder if, like the mean growth rate, these traits can in principle be subject to natural selection. Indeed, a simulation indicates that mutants with more noisy DT distributions, but the same mean DT, can reach fixation in populations of wild-type cells (Supplemental Experimental Procedures). These results show that these mutants can overcome the drift barrier to reach fixation at similar frequencies as a mutant that has the same fitness advantage by having shorter mean DTs, but the same DT noise. Further, for natural selection to act directly on the phenotype of growth variability, it would have to be a genetically-encoded trait. Our results suggest that there are indeed significant genetic determinants that shape the level of growth noise (Figure 6; [9]). We have shown that noisy single-cell growth on palatinose can be lowered by the overexpression of positive feedback in a catabolic gene circuit (Figure 6 and [10]), with inconsistent effects on the mean. Together, this suggest that the trait of increased DT variation and epigenetic inheritance can in principle rise in frequency via natural selection. That said, it seems likely that selection mainly acts on the mean DT, with the effect of noise at best playing a less prominent role.

Finally, our results and model should be more generally useful for predicting the growth rates of populations consisting of different subgroups that have variable growth rates, where clonal variability and/or epigenetic effects are strong [3, 49–53]. For example, tumor growth is shaped by heterogeneous single-cell growth behaviors, however the degree to which these traits affect population-level growth has not been quantitatively described. Our findings therefore provide a framework for more detailed models of cellular growth and help explain how individual-level

variability affects population growth dynamics in diseases that involve clonal growth, such as cancer and microbial pathogenesis.

Experimental Procedures

Yeast strains and growth media used

For the time-lapse microscopy experiments, a genetically diverse set of yeast strains was used. We have performed 12 experiments using the lab strains BY/S288c and SK1 [9, 54, 55], 14 experiments using homozygous diploid strains derived from wild isolates (YPS606, L-1374, Y12, DBVPG1106, DBVPG6765, Y55, YPS128 and BC187; see Liti *et al*., 2009) and 15 experiments using experimentally evolved isolates [9]. For the data shown in Figure 6, a set of 3 strains that are not included in the dataset were used (derived from BY/S288c). For construction details, refer to Supplemental Experimental Information. All experiments were performed at 30°C using YP-based media. The media that were used were YP supplemented with 3% and 10% glucose, 20% maltose, 2.5% galactose, 2% palatinose, 2% glycerol and 2% glucose + 5% ethanol. For more details on the specific combinations of strains and media used in the experiments, see Dataset S1.

Acquisition of time-lapse movies

All strains were first pre-grown in liquid culture at low cell densities ($<2*10^6$ cells per mL) by serially passaging them to achieve balanced steady-state populations (minimally 16h). These cultures were then sandwiched between an agar gel containing the appropriate medium and a coverslip, allowing us to track hundreds of single cells at various positions on the agar pad in a single experiment by periodically taking differential interference contrast (DIC) microscopic images [9]. Also see Supplemental Experimental Information.

Analysis of single-cell DTs

In concordance with previous studies [43], cells were scored either as mothers or daughters. By definition, all cells are born as daughters and become mothers only after finishing their first cell

division. For details on how these DTs were scored, see Figure 1A, Movie S1 and Supplemental Experimental Information.

Predicting population growth rates using the stochastic model

Population growth rates under the stochastic model are predicted using Monte Carlo simulations (see Supplemental Experimental Information). Also see the Results section and Supplemental Information for information on the theoretical framework behind the model, and the derivation of mathematical equations used in the model.

Predicting population growth rates using the deterministic model

We have calculated predicted population growth using the classic model by [43] using the following equation: $e^{(-popGR \times \mu_{mother})} + e^{(-popGR \times \mu_{daughter})} = 1$, with popGR as population growth rate, μ_{mother} as mean mother DT and $\mu_{daughter}$ as mean daughter DT (equation 8 in [43]).

Author Contributions

AMN, BC and KJV conceived and designed the study. BC and AMN performed the experiments and analyzed the results. KP inspired investigating the link between single-cell growth variability and gene expression in palatinose (Figure 6) and constructed the necessary strains. BC conceived the models and wrote the scripts for its simulations. BC and AMN performed statistical analyses of the data. AMN, BC and KJV wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

BC receives a doctoral fellowship grant (Aspirant) from Research Foundation Flanders (FWO-Vlaanderen). AMN acknowledges support from the AB-InBev Baillet-Latour Foundation, Marie Curie Actions and an EMBO Long-term fellowship, ALTF-505-2014. KP is a beneficiary of a mobility grant from the Belgian Federal Science Policy Office. KJV acknowledges funding from the European Union's Horizon 2020 programme (ERC Consolidator Grant CoG682009), HFSP program grant RGP0050/2013,

KU Leuven NATAR Program Financing, VIB, EMBO YIP program, FWO, and IWT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest

The authors have declared that no competing interests exist.

References

1. Stearns, S. (1976). Life-history tactics: a review of the ideas. Q. Rev. Biol. *51*, 3–47.
2. McAdams, H. H., and Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. Trends Genet. *15*, 65–9.
3. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. Science *305*, 1622–1625.
4. Maheshri, N., and O'Shea, E. K. (2007). Living with noisy genes: how cells function reliably with inherent variability in gene expression. Annu. Rev. Biophys. Biomol. Struct. *36*, 413–34.
5. Satory, D., Gordon, A. J. E., Halliday, J. A., and Herman, C. (2011). Epigenetic switches: can infidelity govern fate in microbes? Curr. Opin. Microbiol. *14*, 212–217.
6. Munsky, B., Neuert, G., and van Oudenaarden, A. (2012). Using gene expression noise to understand gene regulation. Science *336*, 183–7.
7. Levy, S. F., Ziv, N., and Siegal, M. L. (2012). Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. PLoS Biol. *10*, e1001325.
8. Avraham, N., Soifer, I., Carmi, M., and Barkai, N. (2013). Increasing population growth by asymmetric segregation of a limiting resource during cell division. Mol. Syst. Biol. *9*, 1–10.
9. New, A. M., Cerulus, B., Govers, S. K., Perez-Samper, G., Zhu, B., Boogmans, S., Xavier, J. B., and Verstrepen, K. J. (2014). Different levels of catabolite repression optimize growth in stable and variable environments. PLoS Biol. *1*, e1001764.
10. Kiviet, D. J., Nghe, P., Walker, N., Boulineau, S., Sunderlikova, V., and Tans, S. J. (2014). Stochasticity of metabolism and growth at the single-cell level. Nature *514*, 376–379.
11. Mars, R. A., Nicolas, P., Ciccolini, M., Reilman, E., Reder, A., Schaffer, M., Mäder, U., Völker, U., van Dijl, J. M., and Denham, E. L. (2015). Small regulatory RNA-induced growth rate heterogeneity of *Bacillus subtilis*. PLoS Genet. *11*, e1005046.
12. Keren, L., van Dijk, D., Weingarten-Gabbay, S., Davidi, D., Jona, G., Weinberger, A., Milo, R., and Segal, E. (2015). Noise in gene expression is coupled to growth rate. Genome Res. *25*, 1893–1902.
13. Kelly, C. D., and Rahn, O. (1932). The growth rate of individual bacterial cells. J. Bacteriol. *23*, 147.
14. Wheals, A., and Lord, P. (1992). Clonal heterogeneity in specific growth rate of *Saccharomyces cerevisiae* cells. Cell Prolif. *25*, 217–23.
15. Wakamoto, Y., Ramsden, J., and Yasuda, K. (2005). Single-cell growth and division dynamics showing epigenetic correlations. Analyst *130*, 311–7.
16. Pin, C., and Baranyi, J. (2006). Kinetics of single cells: observation and modeling of a stochastic process. Appl. Environ. Microbiol. *72*, 2163–9.
17. Di Talia, S., Skotheim, J. M., Bean, J. M., Siggia, E. D., and Cross, F. R. (2007). The effects of molecular noise and size control on variability in the budding yeast cell cycle. Nature *448*, 947–51.
18. Reshes, G., Vanounou, S., Fishov, I., and Feingold, M. (2008). Timing the start of division in *E. coli*: a single-cell study. Phys. Biol. *5*, 046001.
19. Aldridge, B. B., Fernandez-suarez, M., Heller, D., Ambravaneswaran, V., Irimia, D., Toner, M., and Fortune, S. M. (2012). Asymmetry and aging of mycobacterial cells lead to variable growth and antibiotic susceptibility. Science *335*, 100–104.
20. Marusyk, A., and Polyak, K. (2010). Tumor heterogeneity: causes and consequences. Biochim Biophys Acta. *1805*, 1–28.
21. Heng, H. H. Q., Bremer, S. W., Stevens, J., Ye, K. J., Miller, F., Liu, G., and Ye, C. J. (2006). Cancer progression by non-clonal chromosome aberrations. J. Cell. Biochem. *98*, 1424–35.
22. Celli, J. P., Rizvi, I., Evans, C. L., Abu-Yousif, A. O., and Hasan, T. (2010). Quantitative imaging reveals heterogeneous growth dynamics and treatment-dependent residual tumor distributions in a three-

dimensional ovarian cancer model. J. Biomed. Opt. *15*, 051603.

23. Raser, J. M., and O'Shea, E. K. (2005). Noise in gene expression: origins, consequences, and control. Science *309*, 2010–3.

24. Ziv, N., Siegal, M. L., and Gresham, D. (2013). Genetic and non-genetic determinants of cell-growth variation assessed by high-throughput microscopy. Mol Biol Evol *30*, 2568–2578.

25. Tănase-Nicola, S., and ten Wolde, P. R. (2008). Regulatory control and the costs and benefits of biochemical noise. PLoS Comput. Biol. *4*, e1000125.

26. Carey, L. B., van Dijk, D., Sloot, P. M., Kaandorp, J. A., and Segal, E. (2013). Promoter sequence determines the relationship between expression level and noise. PLoS Biol. *11*, e1001528.

27. Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183–6.

28. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. Nat. Genet. *31*, 69–73.

29. Raser, J. M., and O'Shea, E. K. (2004). Control of stochasticity in eukaryotic gene expression. Science *304*, 1811–4.

30. Burga, A., Casanueva, M. O., and Lehner, B. (2011). Predicting mutation outcome from early stochastic variation in genetic interaction partners. Nature *480*, 250–3.

31. Raj, A., Rifkin, S. A., Andersen, E., and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. Nature *463*, 913–8.

32. Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. Cell *135*, 216–26.

33. Fraser, H. B., Hirsh, A. E., Giaever, G., Kumm, J., and Eisen, M. B. (2004). Noise minimization in eukaryotic gene expression. PLoS Biol. *2*, e137.

34. Acar, M., Mettetal, J. T., and Van Oudenaarden, A. (2008). Stochastic switching as a survival strategy in fluctuating environments. Nat. Genet. *40*, 471–475.

35. Beaumont, H. J. E., Gallie, J., Kost, C., Ferguson, G. C., and Rainey, P. B. (2009). Experimental evolution of bet hedging. Nature *462*, 90–93.

36. Casanueva, M. O., Burga, A., and Lehner, B. (2012). Fitness trade-offs and environmentally induced mutation buffering in isogenic *C. elegans*. Science *335*, 82–5.

37. Fridman, O., Goldberg, A., Ronin, I., Shoresh, N., and Balaban, N. Q. (2014). Optimization of lag time underlies antibiotic tolerance in evolved bacterial populations. Nature *513*, 418–421.

38. Zhang, Z., Qian, W., and Zhang, J. (2009). Positive selection for elevated gene expression noise in yeast. Mol. Syst. Biol. *5*, 299.

39. Wang, Z., and Zhang, J. (2011). Impact of gene expression noise on organismal fitness and the efficacy of natural selection. Proc. Natl. Acad. Sci. U. S. A. *108*, E67–76.

40. Metzger, B. P. H., Yuan, D. C., Gruber, J. D., Duveau, F., and Wittkopp, P. J. (2015). Selection on noise constrains variation in a eukaryotic promoter. Nature *521*, 344–347.

41. Siegal-Gaskins, D., and Crosson, S. (2008). Tightly regulated and heritable division control in single bacterial cells. Biophys. J. *95*, 2063–72.

42. Sandler, O., Mizrahi, S. P., Weiss, N., Agam, O., Simon, I., and Balaban, N. Q. (2015). Lineage correlations of single cell division time as a probe of cell-cycle dynamics. Nature *519*, 468–471.

43. Hartwell, L. H., and Unger, M. W. (1977). Unequal division in *Saccharomyces cerevisiae* and its implications for the control of cell division. J. Cell Biol. *75*, 422–435.

44. Lord, P. G., and Wheals, A. E. (1980). Asymmetrical division of *Saccharomyces cerevisiae*. J. Bacteriol. *142*, 808–18.

45. Moore, L. S., Stolovicki, E., and Braun, E. (2013). Population dynamics of metastable growth-rate phenotypes. PLoS One *8*, e81671.

46. Brown, C., Murray, A. W., and Verstrepen, K. J. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. Curr. Biol. *20*, 895–903.

47. Teste, M.-A., François, J. M., and Parrou, J.-L. (2010). Characterization of a new multigene family encoding isomaltases in the yeast *Saccharomyces cerevisiae*, the IMA family. J. Biol. Chem. *285*, 26815–24.

48. Voordeckers, K., Brown, C. A., Vanneste, K., van der Zande, E., Voet, A., Maere, S., and Verstrepen, K. J. (2012). Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. PLoS Biol. *10*, e1001446.

49. Helaine, S., and Holden, D. W. (2013). Heterogeneity of intracellular replication of bacterial pathogens. Curr. Opin. Microbiol. *16*, 184–91.

50. Roesch, A., and Fukunaga-Kalabis, M. (2010). A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell *141*, 583–594.

51. Huang, S., and Ingber, D. E. (2007). A non-genetic basis for cancer progression and metastasis: self-organizing attractors in cell regulatory networks. Breast Dis. *26*, 27–54.

52. Kaneko, K. (2011). Characterization of stem cells and cancer cells on the basis of gene expression profile stability, plasticity, and robustness: dynamical systems theory of gene expressions under cell-cell interaction explains mutational robustness of differentiated cells. Bioessays *33*, 403–13.

53. Gupta, P. B., Fillmore, C. M., Jiang, G., Shapira, S. D., Tao, K., Kuperwasser, C., and Lander, E. S. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell *146*, 633–44.

54. Brachmann, C., Davies, A., Cost, G., Caputo, E., Li, J., Hieter, P., and Boeke, J. (1998). Deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast *14*, 115–132.

55. Liti, G., Carter, D. M., Moses, A. M., Warringer, J., Parts, L., James, S. A., Davey, R. P., Roberts, I. N., Burt, A., Koufopanou, V., et al. (2009). Population genomics of domestic and wild yeasts. Nature *458*, 337–341.

Figure legends

Figure 1. Measurement of single-cell division times of mother and daughter cells using time-lapse microscopy.

(A) A time-lapse microscopy image series of *S. cerevisiae* strain BY/S288c growing in YP + 3% glucose is shown to indicate how cell division events (indicated by the arrows) were scored. DTs are determined by the time difference between consecutive cell divisions. For daughter cells, the DT is divided into an unbudded period and a budded period, based on the appearance of the first bud. See also Figure S1A and Movie S1.

(B) Mother DT (left) and daughter DT (right) distributions for 10 genetically different yeast strains in up to 7 different growth media (N=41 experiments). The distributions highlighted represent Y55 (green) and BC187 (purple) growing in 2.5% galactose.

Figure 2. Architecture of individual yeast DT traits.

(A) Mean mother and daughter DTs are linearly correlated across different yeast strains growing exponentially in different carbon sources (Dataset S1).

(B) The length of the unbudded and budded period is plotted against the total daughter DT, with each data point representing a single-cell measurement (data comprise all usable daughter DT measurements obtained in this study, n=2547 cells).

(C) DT coefficient of variation (CoV) is plotted against mean mother (blue) and daughter (red) DT. The green and purple fills (respectively the strains Y55 and BC187, growing in 2.5% galactose) illustrate how some strains can have the same mean DT but different variance.

(D) The DT length is epigenetically inherited across closely related cells within a lineage. This graph shows the correlation between DTs of a mother and her most recently born daughter (z-scores of $mDT_{1/2}$ vs. z-scores of $dDT_{1/2}$) across the whole dataset. See also Figure S1B-G.

All error bars represent bootstrapped SDs (Supplemental Experimental Procedures).

Figure 3. An individual-based model to predict population growth rate.

(A) All cells are born as daughter cells (originating from the dashed arrows), and become mothers after completion of their first bud. At birth or after the completion of a bud, each cell is assigned a new DT that is dependent on the DT variance and epigenetic DT behaviors. The model can also be run without taking into account epigenetic behaviors ($model_{stoch,no\text{-}epig}$) or stochasticity ($model_{determ}$). Also see Results section and Figure S2.

(B) Modeled population growth rates ($model_{stoch,no\text{-}epig}$) vs. growth rates measured on solid media (microcolony growth rates). Microcolony growth rates were measured by microscopically tracking the cell count increase of microcolonies over time.

(C) Modeled population growth rates ($model_{stoch,no\text{-}epig}$) vs. growth rates measured in liquid cultures. Also see Figure S3.

Error bars represent bootstrapped SDs (Supplemental Experimental Procedures).

Figure 4. Heterogeneous single-cell growth affects population-level growth rates.

(A) The ratio of predicted population growth rates of $model_{stoch,no\text{-}epig}$ over the deterministic model is plotted against the daughter DT CoV for all experiments in the dataset. Error bars represent bootstrapped SDs (Supplemental Experimental Procedures). Similar results are observed when using Gamma and Pearson system distributions or the empirical distributions (Figure S4).

(B) The predicted population growth rate depends on the variability within the mother and daughter DT distributions, and the level of mother-daughter mean asymmetry. At each level of asymmetry, the mother DT CoV and daughter DT CoV were varied independently and the ratios of the predicted population growth rates of $model_{stoch,no\text{-}epig}$ over the deterministic model were plotted as heat maps. The black dots represent the CoV values observed in our dataset plotted on the heat map with the closest corresponding mother-daughter asymmetry.

Figure 5. Epigenetic inheritance of single-cell growth rates further increases population-level growth rates.

(A) Two experimentally observed DT correlations were considered in the model: (up) mDT_1 vs. mDT_2 and (down) $mDT_{1/2}$ vs. $dDT_{1/2}$. Here, histograms of the R^2-values of these correlations are shown per experiment. The asterisk marks the bin that includes BY/S288c in 2% palatinose.

(B) The ratio of the growth rates of the model with epigenetics ($model_{stoch,epig}$) and without epigenetics ($model_{stoch,no\text{-}epig}$) were calculated for each experiment and sorted from low to

high. The asterisk marks the data point corresponding to BY/S288c in 2% palatinose. See also Figure S5.

(C) The effect of epigenetic single-cell DT correlations on the predicted population-level growth rate is illustrated with a heat map. Colors are fold-increase of a given model over the deterministic model. The dots (I-III) are predictions of the deterministic model (I) or the stochastic model without (II) or with (III) epigenetics for BY/S288 growing in 2% palatinose.

Figure 6. Overexpressing the genes required for growth reduces the DT variability and epigenetic DT inheritance.

(A) Palatinose is transported into the cytoplasm by Mal11p, and broken down by Ima1p. In the presence of intracellular palatinose, *MAL11* and *IMA1* expression is induced by two transcriptional activators (represented here by PalR), a typical positive feedback motif.

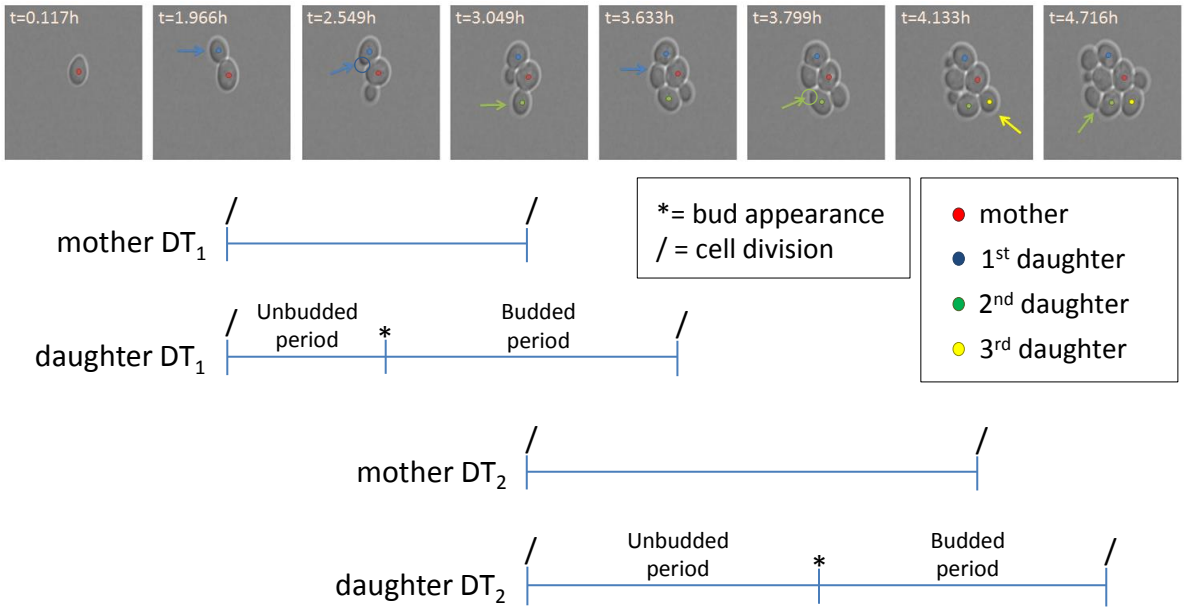
(B) Flow cytometry analysis of the effect of MAL11 overexpression. Shown are traces of IMA1 expression of the wild-type BY/S288c (red); a strain in which *MAL11* and *IMA1* have been fluorescently tagged (blue); and the dually-tagged fluorescent strain in which *MAL11* has been overexpressed (green). Gene expression noise is indicated as $\frac{\sigma^2}{\mu^2}$ [26].

(c) Epigenetic DT inheritance in the wild-type and a *MAL11* overexpression strain. In this plot, jitter was added in both dimensions to represent better the density of the observations.

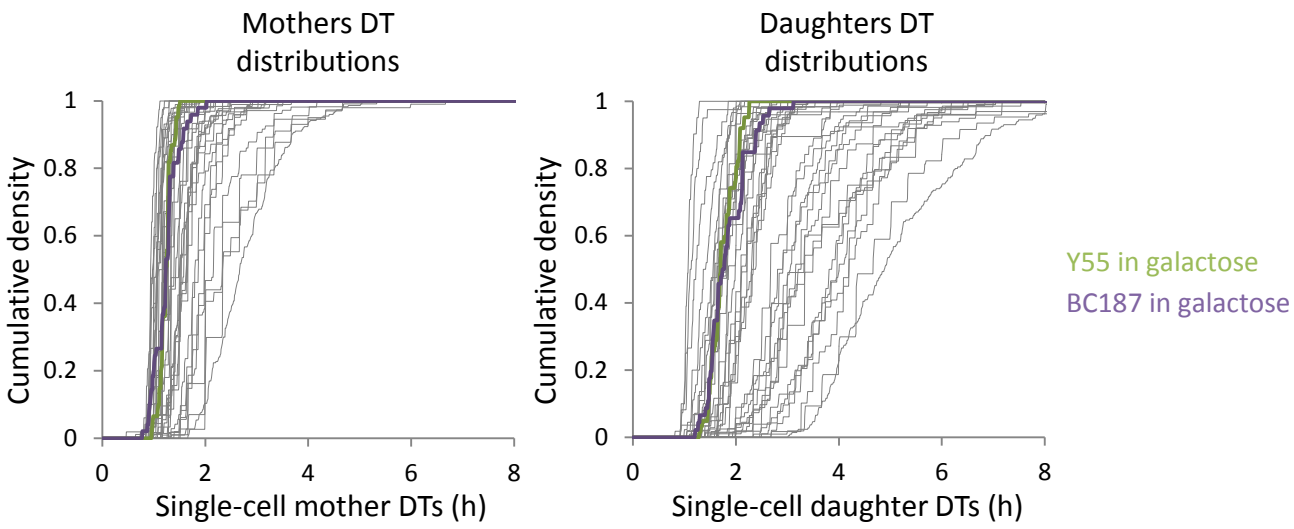
Supplemental Information

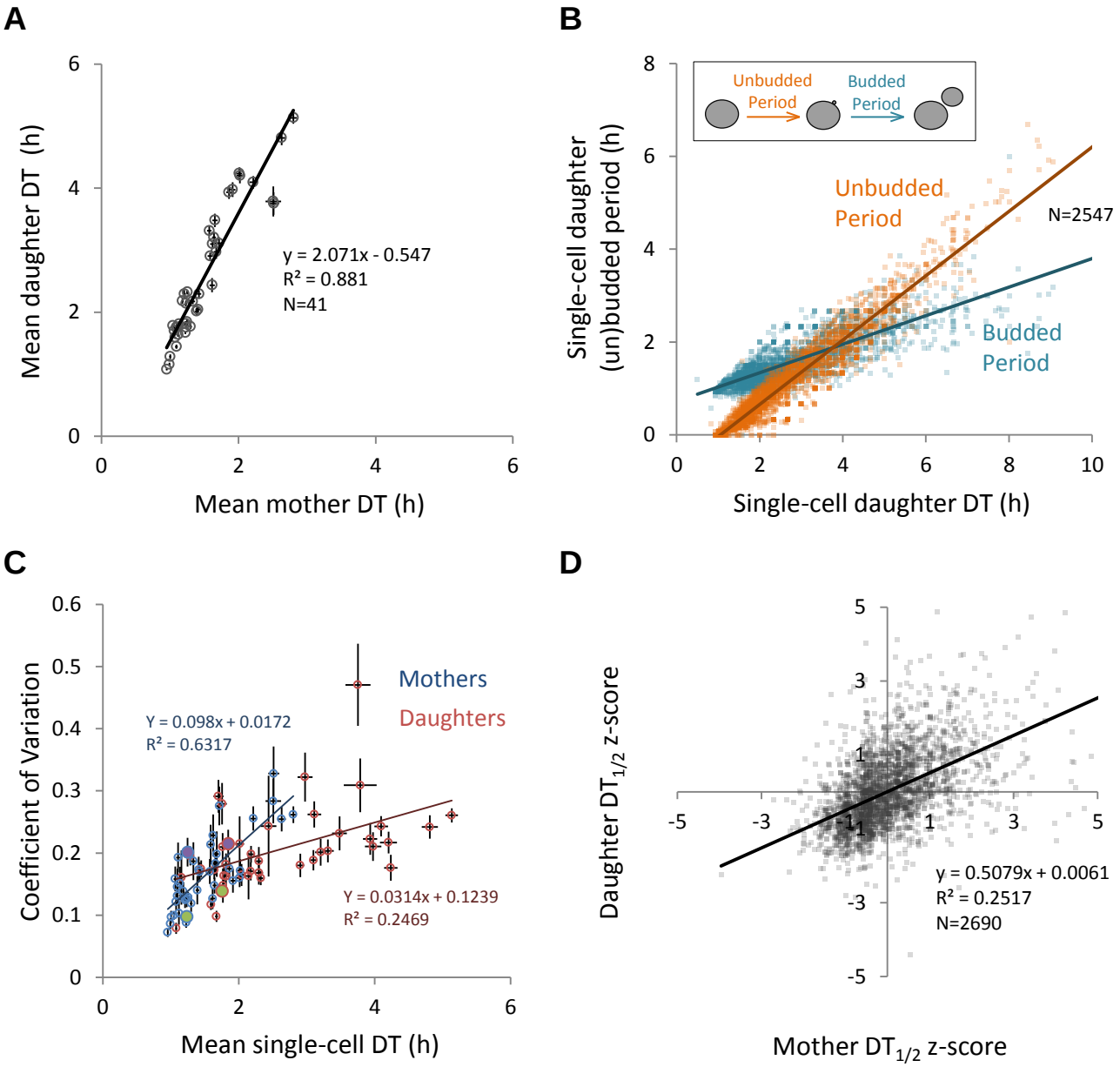
Supplemental Information. Contains the Supplemental Experimental Procedures and Figures S1-S6.

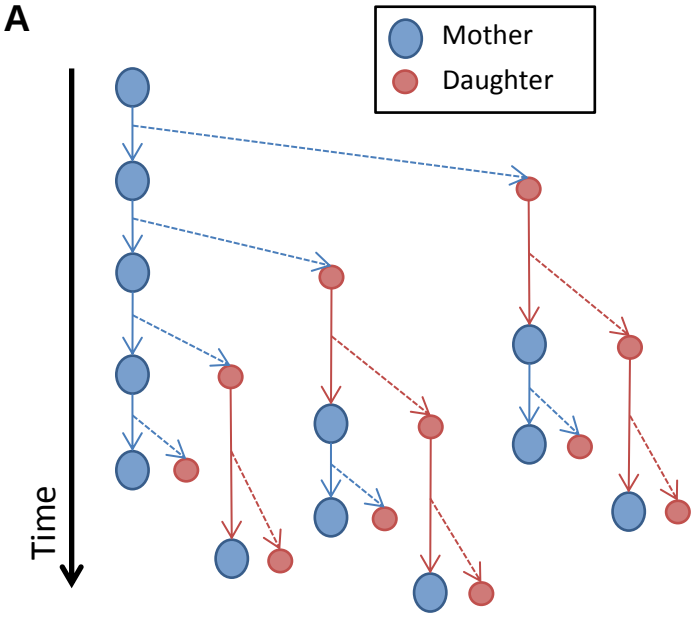
A


t=0.117h
t=1.966h
t=2.549h
t=3.049h
t=3.633h
t=3.799h
t=4.133h
t=4.716h
mother DT1
daughter DT1
Unbudded period
Budded period
*= bud appearance
/ = cell division
mother
1st daughter
2nd daughter
3rd daughter
mother DT2
daughter DT2
Unbudded period
Budded period

B


Mothers DT distributions
Daughters DT distributions
Cumulative density
Single-cell mother DTs (h)
Single-cell daughter DTs (h)
Y55 in galactose
BC187 in galactose


A
Mean daughter DT (h)
Mean mother DT (h)
y = 2.071x - 0.547
R² = 0.881
N=41
B
Single-cell daughter (un)budded period (h)
Single-cell daughter DT (h)
Unbudded Period
Budded Period
Unbudded Period
Budded Period
N=2547
C
Coefficient of Variation
Mean single-cell DT (h)
Mothers
Daughters
Y = 0.098x + 0.0172
R² = 0.6317
Y = 0.0314x + 0.1239
R² = 0.2469
D
Daughter DT1/2 z-score
Mother DT1/2 z-score
y = 0.5079x + 0.0061
R² = 0.2517
N=2690



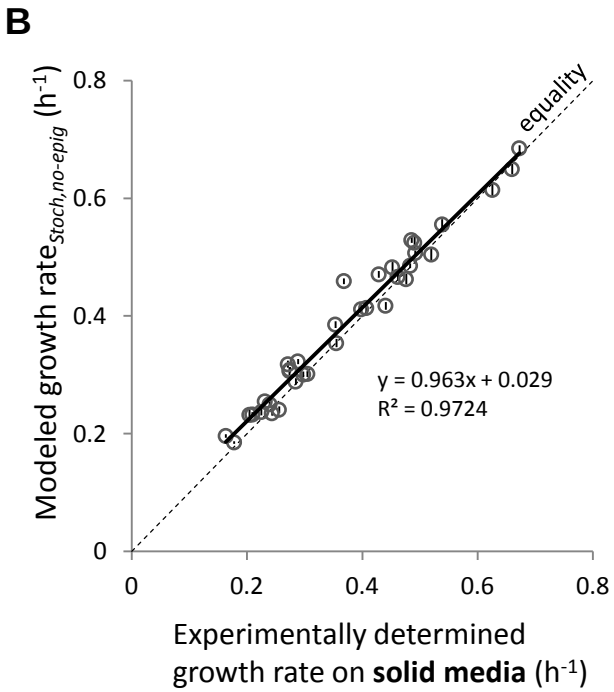
A model predicting population growth from single-cell DTs

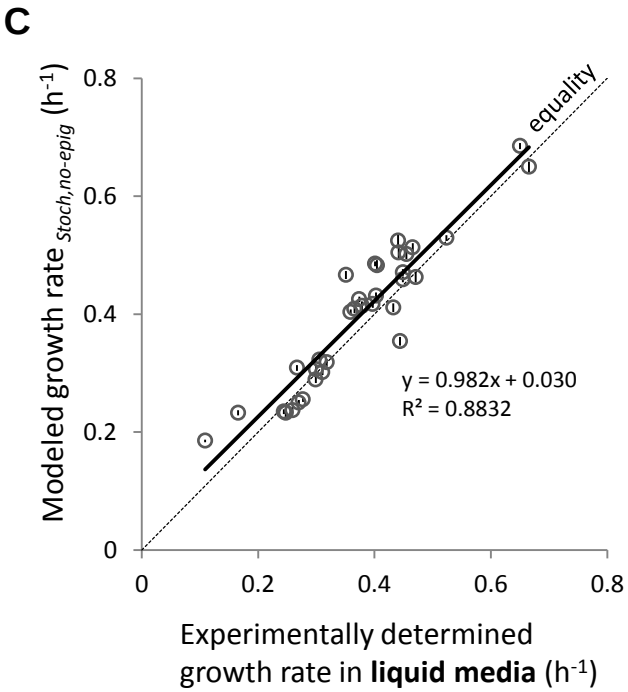
$Model_{Stoch,epig}$ | ***Including DT variability and epigenetic behavior***

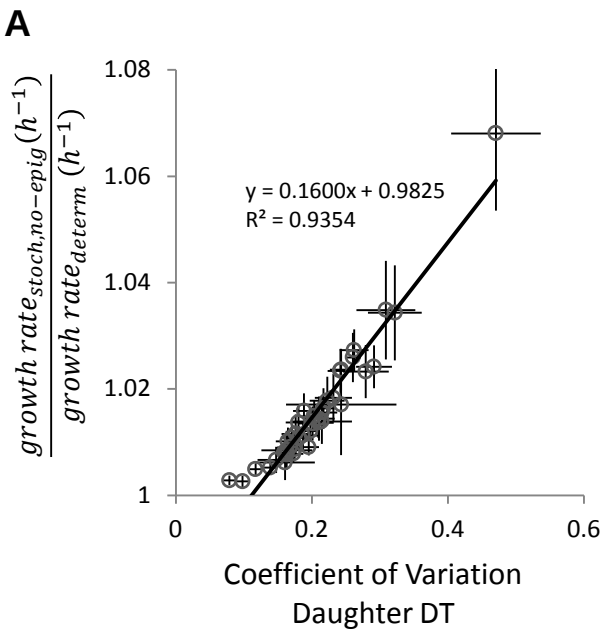
- Mother DT = f (μ_{mother}, σ_{mother}^2, R^2 (mDT_1 vs. mDT_2))
- Daughter DT = f ($\mu_{daughter}$, $\sigma_{daughter}^2$, R^2($mDT_{1/2}$ vs. $dDT_{1/2}$))

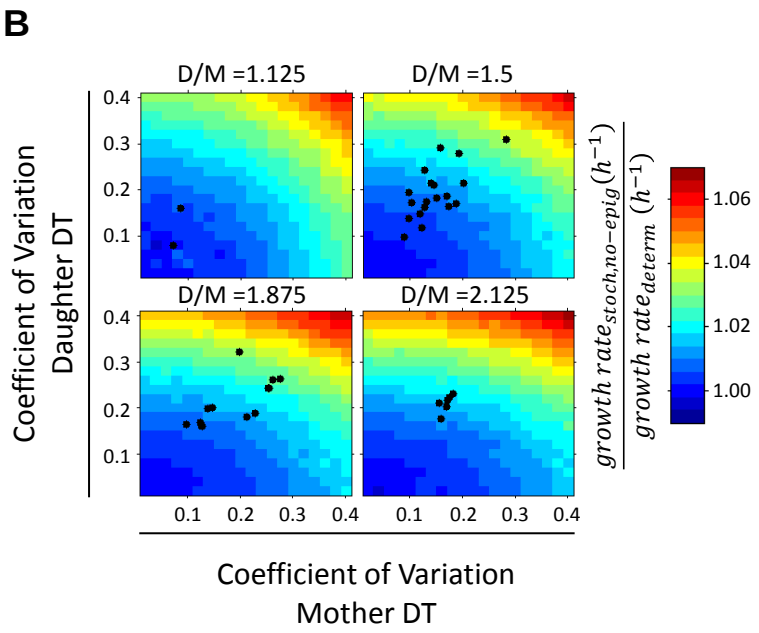
$Model_{Stoch,no-epig}$ | *for a strictly stochastic model*: $R^2 = 0$

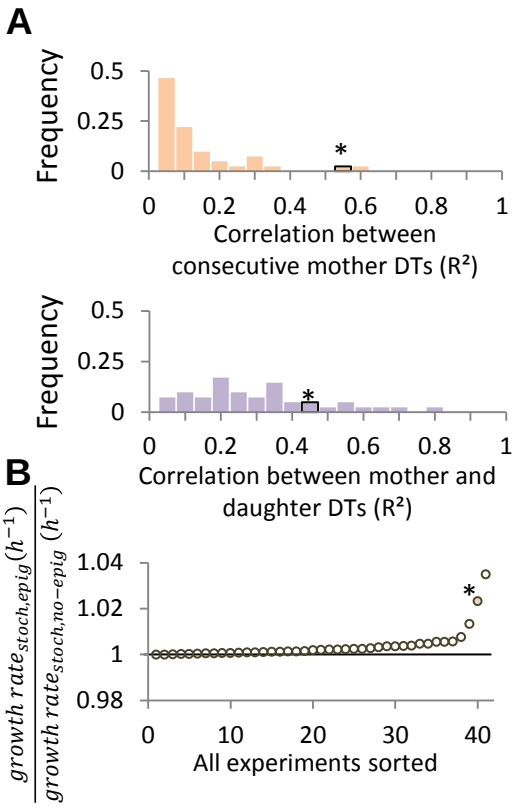
$Model_{Determ}$ | *for strictly deterministic model*: $R^2 = 0$ and $\sigma^2 = 0$

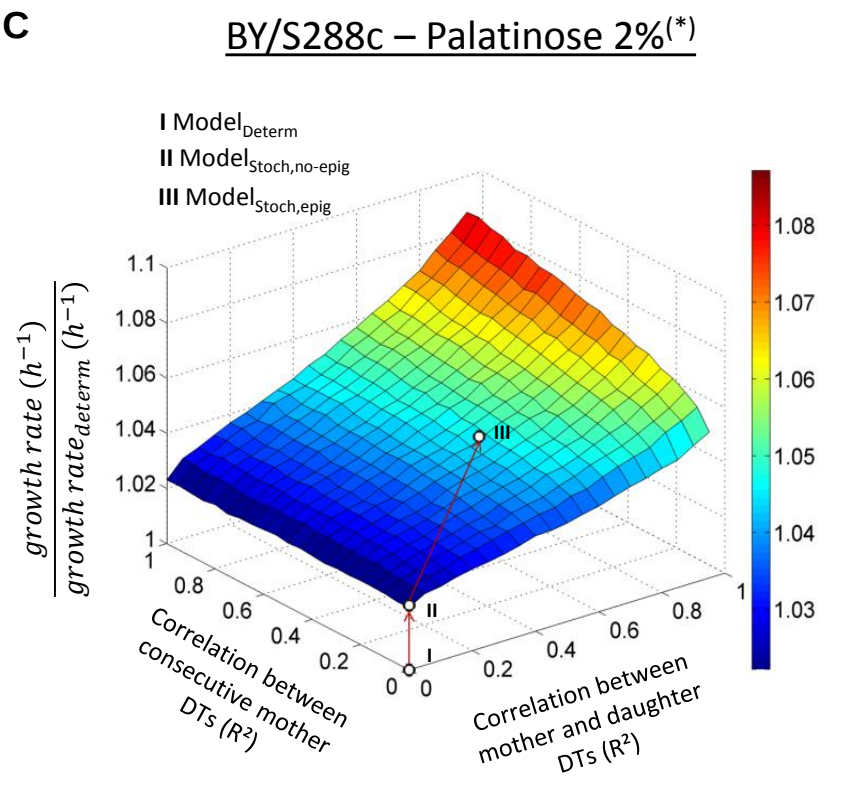





A
1.08
1.06
1.04
1.02
1
0
0.2
0.4
0.6
y = 0.1600x + 0.9825
R² = 0.9354
$\frac{growth\ rate_{stoch,no-epig}(h^{-1})}{growth\ rate_{determ}\ (h^{-1})}$
Coefficient of Variation
Daughter DT


B
D/M =1.125
D/M =1.5
D/M =1.875
D/M =2.125
0.4
0.3
0.2
0.1
0.1
0.2
0.3
0.4
Coefficient of Variation
Daughter DT
Coefficient of Variation
Mother DT
$\frac{growth\ rate_{stoch,no-epig}(h^{-1})}{growth\ rate_{determ}\ (h^{-1})}$
1.06
1.04
1.02
1.00


A
Frequency
0.5
0.25
0
0 0.2 0.4 0.6 0.8 1
Correlation between consecutive mother DTs (R²)
Frequency
0.5
0.25
0
0 0.2 0.4 0.6 0.8 1
Correlation between mother and daughter DTs (R²)
B
$\frac{growth\ rate_{stoch,epig}(h^{-1})}{growth\ rate_{stoch,no-epig}(h^{-1})}$
1.04
1.02
1
0.98
0 10 20 30 40
All experiments sorted


C
BY/S288c – Palatinose 2%(*)
I Model$_{Determ}$
II Model$_{Stoch,no-epig}$
III Model$_{Stoch,epig}$
$\frac{growth\ rate\ (h^{-1})}{growth\ rate_{determ}\ (h^{-1})}$
1.1
1.08
1.06
1.04
1.02
1
III
II
I
Correlation between consecutive mother DTs (R²)
Correlation between mother and daughter DTs (R²)
1.08
1.07
1.06
1.05
1.04
1.03

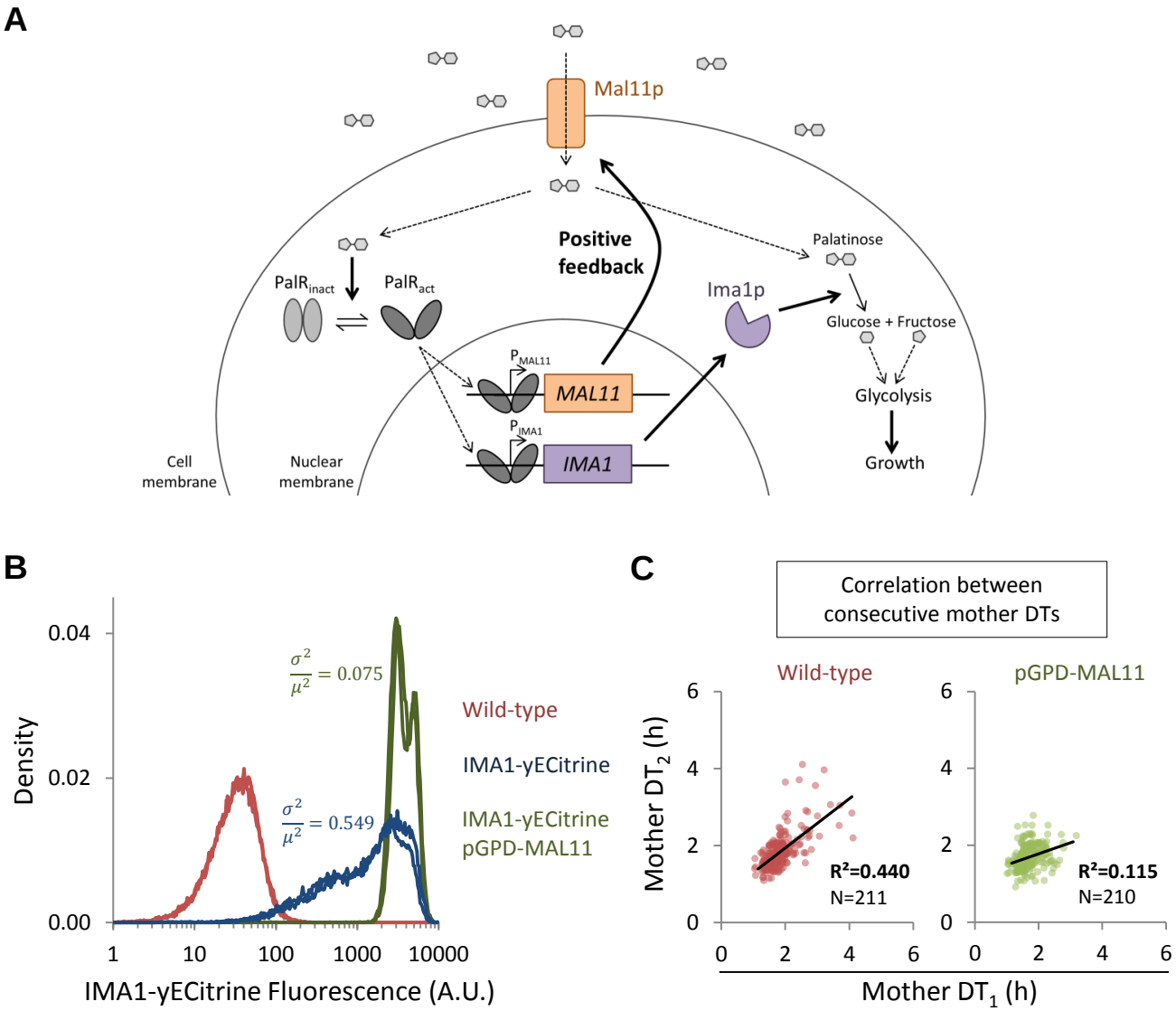

A
Mal11p
Positive feedback
Palatinose
PalR$_{inact}$
PalR$_{act}$
Ima1p
Glucose + Fructose
P$_{MAL11}$
MAL11
Glycolysis
P$_{IMA1}$
IMA1
Growth
Cell membrane
Nuclear membrane
B
0.04
0.02
0.00
Density
$\frac{\sigma^2}{\mu^2} = 0.075$
$\frac{\sigma^2}{\mu^2} = 0.549$
Wild-type
IMA1-yECitrine
IMA1-yECitrine pGPD-MAL11
1
10
100
1000
10000
IMA1-yECitrine Fluorescence (A.U.)
C
Correlation between consecutive mother DTs
Wild-type
pGPD-MAL11
Mother DT$_2$ (h)
6
4
2
0
R²=0.440
N=211
R²=0.115
N=210
0
2
4
6
Mother DT$_1$ (h)

Figure S1

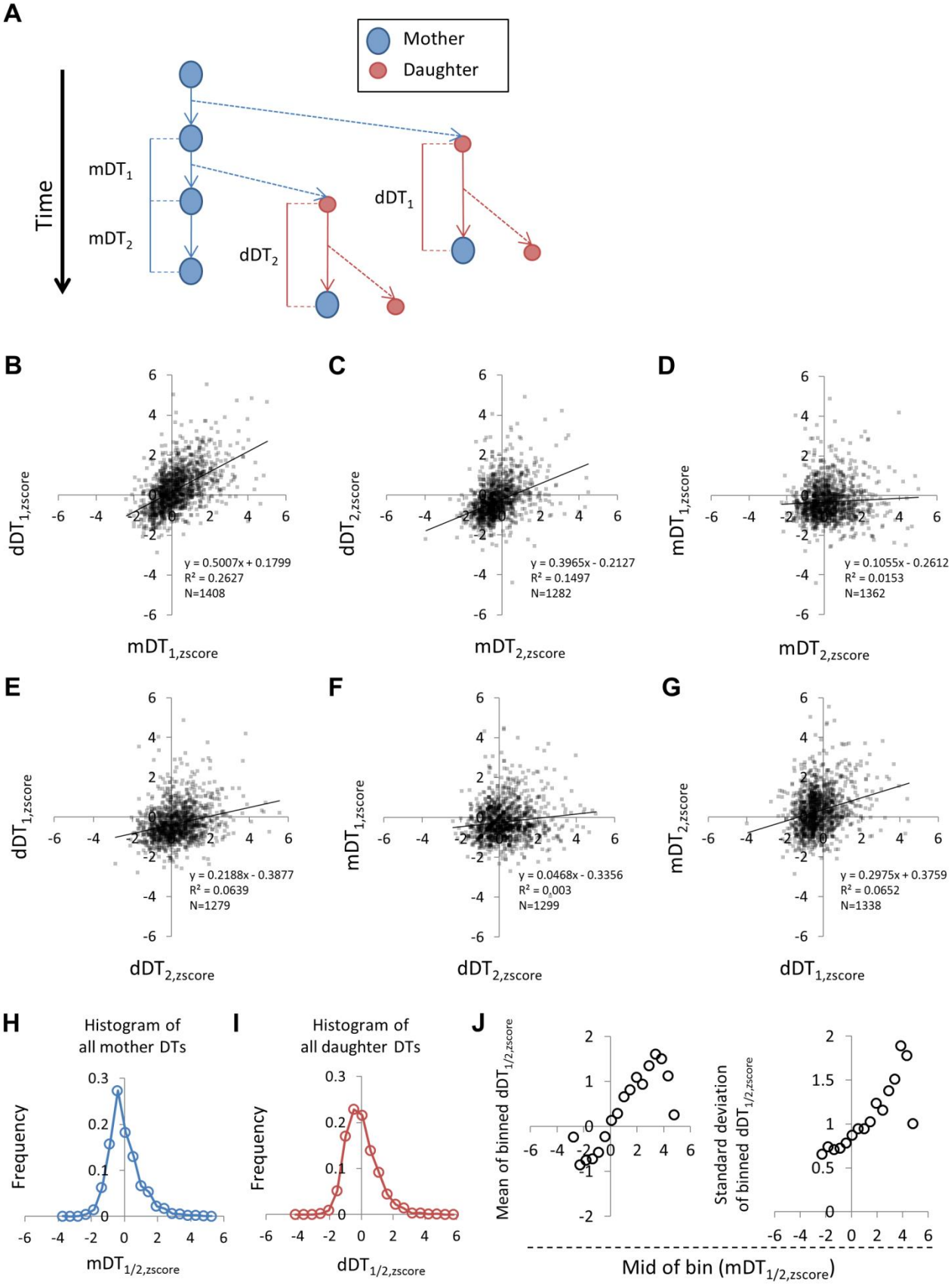

A
Mother
Daughter
Time
mDT1
mDT2
dDT1
dDT2
B
dDT1,zscore
mDT1,zscore
y = 0.5007x + 0.1799
R² = 0.2627
N=1408
C
dDT2,zscore
mDT2,zscore
y = 0.3965x - 0.2127
R² = 0.1497
N=1282
D
mDT1,zscore
mDT2,zscore
y = 0.1055x - 0.2612
R² = 0.0153
N=1362
E
dDT1,zscore
dDT2,zscore
y = 0.2188x - 0.3877
R² = 0.0639
N=1279
F
mDT1,zscore
dDT2,zscore
y = 0.0468x - 0.3356
R² = 0.003
N=1299
G
mDT2,zscore
dDT1,zscore
y = 0.2975x + 0.3759
R² = 0.0652
N=1338
H
Histogram of all mother DTs
Frequency
mDT1/2,zscore
I
Histogram of all daughter DTs
Frequency
dDT1/2,zscore
J
Mean of binned dDT1/2,zscore
Standard deviation of binned dDT1/2,zscore
Mid of bin (mDT1/2,zscore)

Figure S1. Related to Figure 1-2. Average DT correlations between all possible genealogical relationships across the whole dataset. The mother and daughter DT measurements from all 41 experiments were first transformed to within-distribution z-score values. In this way, cells growing faster than the average have negative z-scores, and vice versa.

(A) Genealogical relationships of mother and daughter cells which were tracked within microcolonies using time-lapse microscopy. The DTs which were scored are indicated by accolades, and the annotations are the same as in Figure 1A. Within one lineage (microcolony) starting from one single cell, two consecutive mother DTs, and two daughter DTs are scored. Note that the first cell can start out either as a mother or a daughter cell, but is a mother cell (by definition) when its first DT is scored.

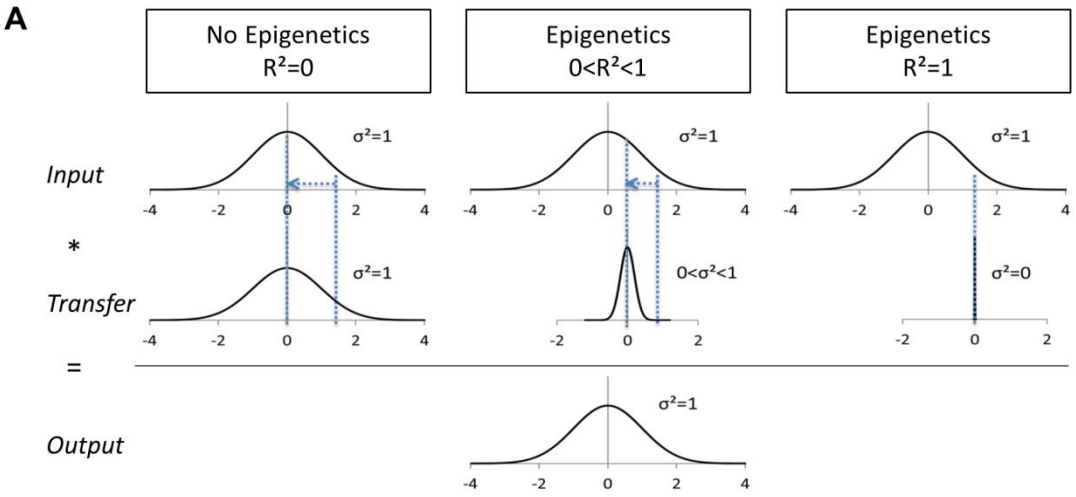
(B-G) The DT z-scores were then grouped per category (mDT_1,mDT_2,dDT_1,dDT_2) without accounting for the different number of DTs scored between the different experiments, and correlations between all possible genealogical relationships were plotted. First, scatter plots of DT z-scores of a mother and her most recently born daughter are shown ((B) $dDT_{1,zscore}$ vs. $mDT_{1,zscore}$ and (C) $dDT_{2,zscore}$ vs. $mDT_{2,zscore}$). Note that these data were combined to produce Figure 2D. Next, the scatter plot of z-scores of consecutive mother DTs is shown in (D). Finally, the remaining correlations are plotted: (E) $dDT_{1,zscore}$ vs. $dDT_{2,zscore}$, (F) $mDT_{1,zscore}$ vs. $dDT_{2,zscore}$ and (G) $mDT_{2,zscore}$ vs. $dDT_{1,zscore}$. The number of data points (N) is respectively 1408, 1282, 1362, 1279, 1299 and 1338 pairs of DTs.

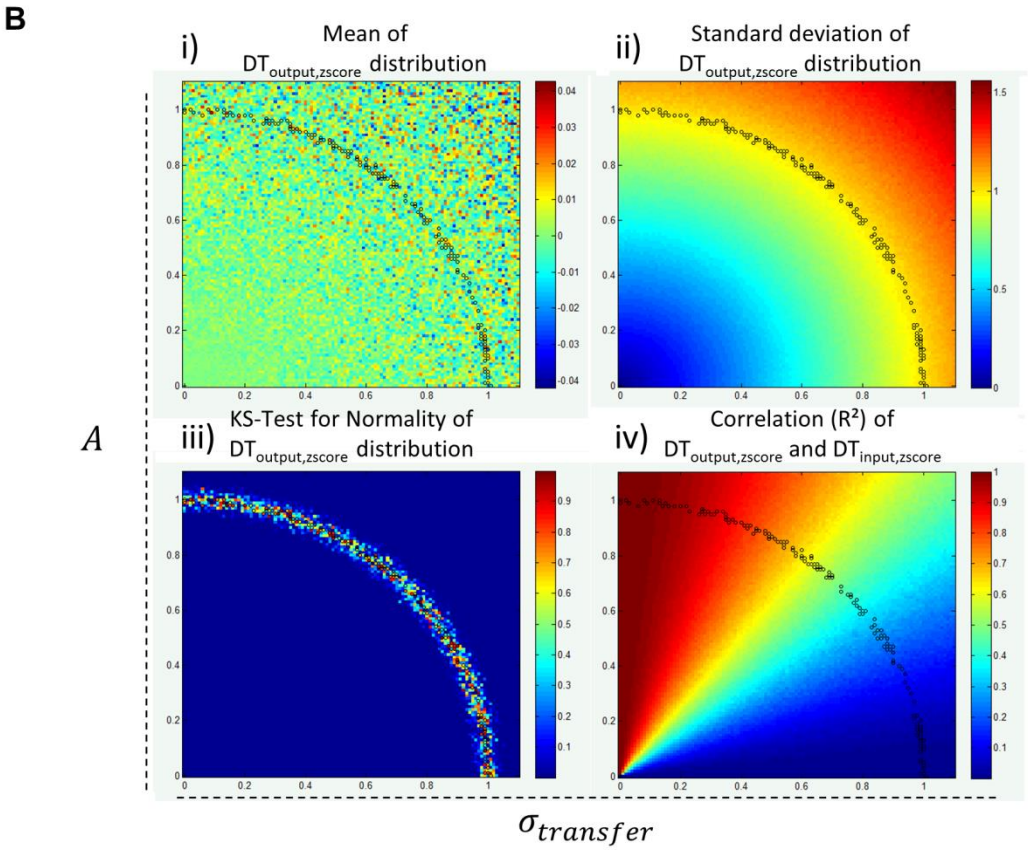
(H) Histogram of all mother DT z-scores represented in Figure 2D (N=2690).

(I) Histogram of all daughter DT z-scores represented in Figure 2D (N=2690).

(J) The data points shown in Figure 2D were binned on the x-axis ($mDT_{1/2,zscore}$) in 19 equally sized bins, and normal distributions were fitted to the resulting y-axis values in each bin ($dDT_{1/2,zscore}$). (left) Mean and (right) SD of these fitted distributions were plotted against the mid of each bin.

Figure S2


A
No Epigenetics
R²=0
Epigenetics
0<R²<1
Epigenetics
R²=1
Input
*
Transfer
=
Output
σ²=1
0<σ²<1
σ²=0


B
A
i) Mean of DT output,zscore distribution
ii) Standard deviation of DT output,zscore distribution
iii) KS-Test for Normality of DT output,zscore distribution
iv) Correlation (R²) of DT output,zscore and DT input,zscore
σ transfer

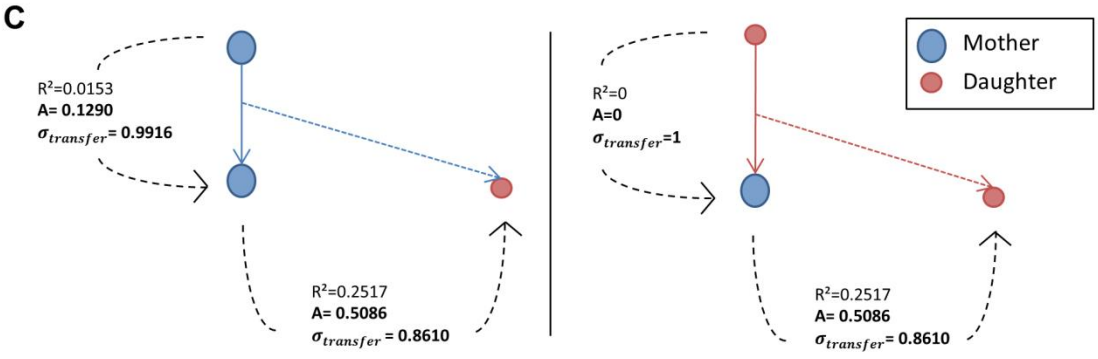

C
R²=0.0153
A= 0.1290
σ transfer= 0.9916
R²=0.2517
A= 0.5086
σ transfer = 0.8610
R²=0
A=0
σ transfer=1
R²=0.2517
A= 0.5086
σ transfer= 0.8610
Mother
Daughter

Figure S2. Related to Figure 3. The stochastic growth model takes into account epigenetic DT inheritance.

(A) Three characteristic scenarios that describe how different extents of correlation (R^2) can arise between two normally distributed DT distributions. This figure gives a visual representation of how new DT are chosen (in)dependent on previously assigned DTs in the stochastic model. In these examples, we start out with an initial set of standard normally distributed values, defined as the input distribution. For each of these values, a new value is chosen individually using a transfer function. In the end, this should result in a new set of standard normally distributed values, defined as the output distribution. (A) No correlation between input and output distributions. (B) An imperfect correlation between input and output distribution, with an R^2 between 0 and 1. (C) A perfect correlation between input and output distribution.

(B) The limiting condition that we impose on the transfer function, is that when provided with a series of normally distributed input DTs, it should return a series of output DTs matching the experimentally observed DT distribution (also see (A)). These plots represent the results of stochastic simulations that show that our transfer function follows these requirements under certain parameter combinations. Furthermore, it allows us to quantitatively determine how combinations of our parameters $\sigma_{transfer}$ and A lead to different extents of correlation (R^2) between the input and output distributions. For each data point represented in the heat maps, a set of 10 000 $DT_{input,zscore}$ sampled from a standard normal distribution were used as an input to acquire respectively 10 000 $DT_{output,zscore}$ using the Supplemental equations 11-13. The values for $\sigma_{transfer}$ and A were systematically varied (from 0 to 1.1 in steps of 0.01) and the outcome evaluated by calculating (i) the mean, (ii) SD and (iii) normality (p-value using Kolmogorov Smirnov-test) of the resulting $DT_{output,zscore}$ distributions. Finally, the correlation (R^2) of the $DT_{output,zscore}$'s against their respective $DT_{input,zscore}$'s is shown in (iv). These four parameters were plotted as heat maps, with the vertical and horizontal axes representing the parameters $\sigma_{transfer}$ and A respectively. The black circles in each panel represent the simulations which led to a mean of 0 (+-0.03) and SD of 1 (+-0.005) in the resulting $DT_{output,zscore}$ distributions.

(C) The epigenetic DT relationships that are incorporated in our model are visualized. The left panel shows the situation of a mother cell dividing, and the right panel of a daughter cell dividing. The black striped arrows show how the z-score of a new DT is chosen. From each newly assigned z-score, the DT is calculated based on the parameters of the normal distributions which were experimentally determined using time-lapse microscopy (Figure 1 and Dataset S2). The parameters $\sigma_{transfer}$ and A were determined for the correlations (R^2) shown in Figure 2D and Figure S1D using Supplemental equations 14 and 15. The specific values of these parameters shown here thus correspond to the average DT correlations across all experiments.

Figure S3

A

Correlation

R

	$Meas_{Solid\text{-}Area}$	$Meas_{Solid\text{-}CellCount}$	$Measured_{Liquid}$
$Model_{Determ}$	0.9807	0.9852	0.9401
$Model_{Stoch,no\text{-}epig}$	0.9813	0.9861	0.9398
$Model_{Stoch,epig}$	0.9817	0.9867	0.9388

Inaccuracy

Mean squared error

$Meas_{Solid\text{-}Area}$	$Meas_{Solid\text{-}CellCount}$	$Measured_{Liquid}$
3.66E-03	6.49E-04	2.02E-03
4.23E-03	7.52E-04	2.23E-03
4.35E-03	7.69E-04	2.30E-03

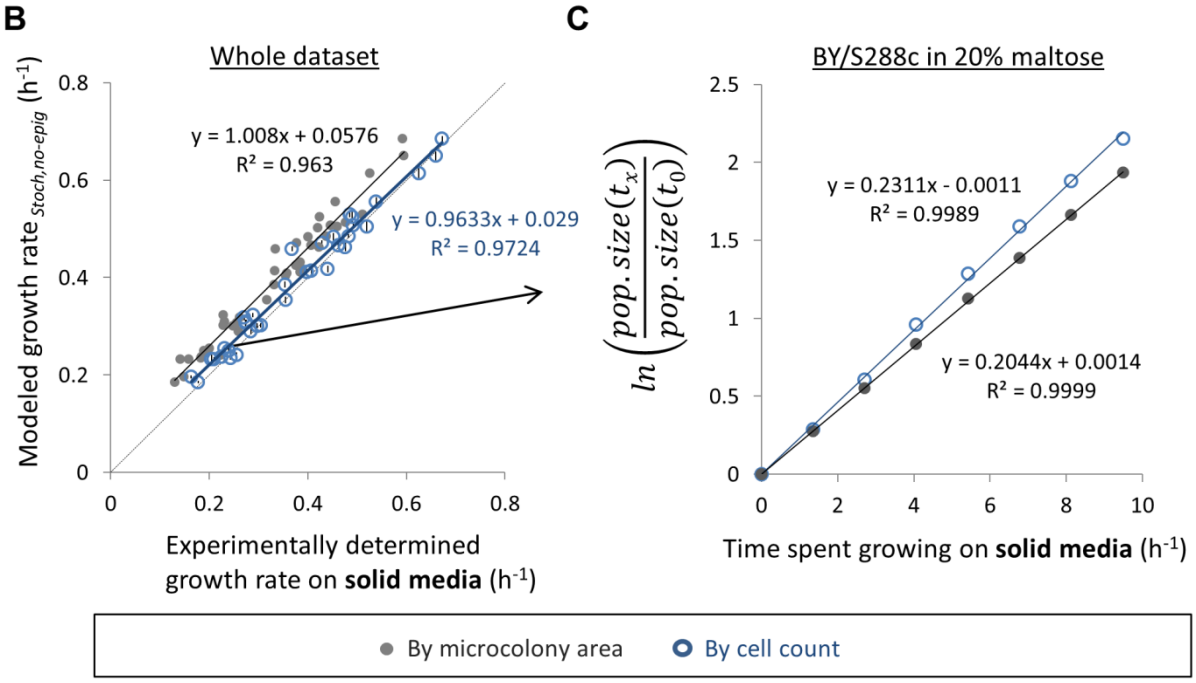


Figure S3. Related to Figure 3. The model's predictions closely match experimental measures of population growth rate.

(A) The predicted population growth rates are analyzed for their ability to predict the measured population growth rates. Growth rates were predicted using the three variants of the model ($model_{determ}$, $model_{stoch,no\text{-}epig}$ and $model_{stoch,epig}$), and measured on solid media (by tracking either the area or cell count increase) and in liquid medium (using colony counting). The correlation was calculated using the Pearson correlation coefficient (R) and the inaccuracy was calculated using the mean squared error.

(B) Modeled population growth rates ($model_{stoch,no\text{-}epig}$) are correlated with growth rates measured on solid media in two different ways. The blue open circles show the same data as in Figure 3B, and represent the correlation with the growth rates measured by tracking the increase in cell count within the microcolonies. The filled grey circles show the correlation with growth rates measured by tracking the increase in microcolony area.

(C) This panel shows how the increase of area and cell count within the same microcolonies for one specific experiment, namely BY/S288c growing in maltose 20%.

Figure S4

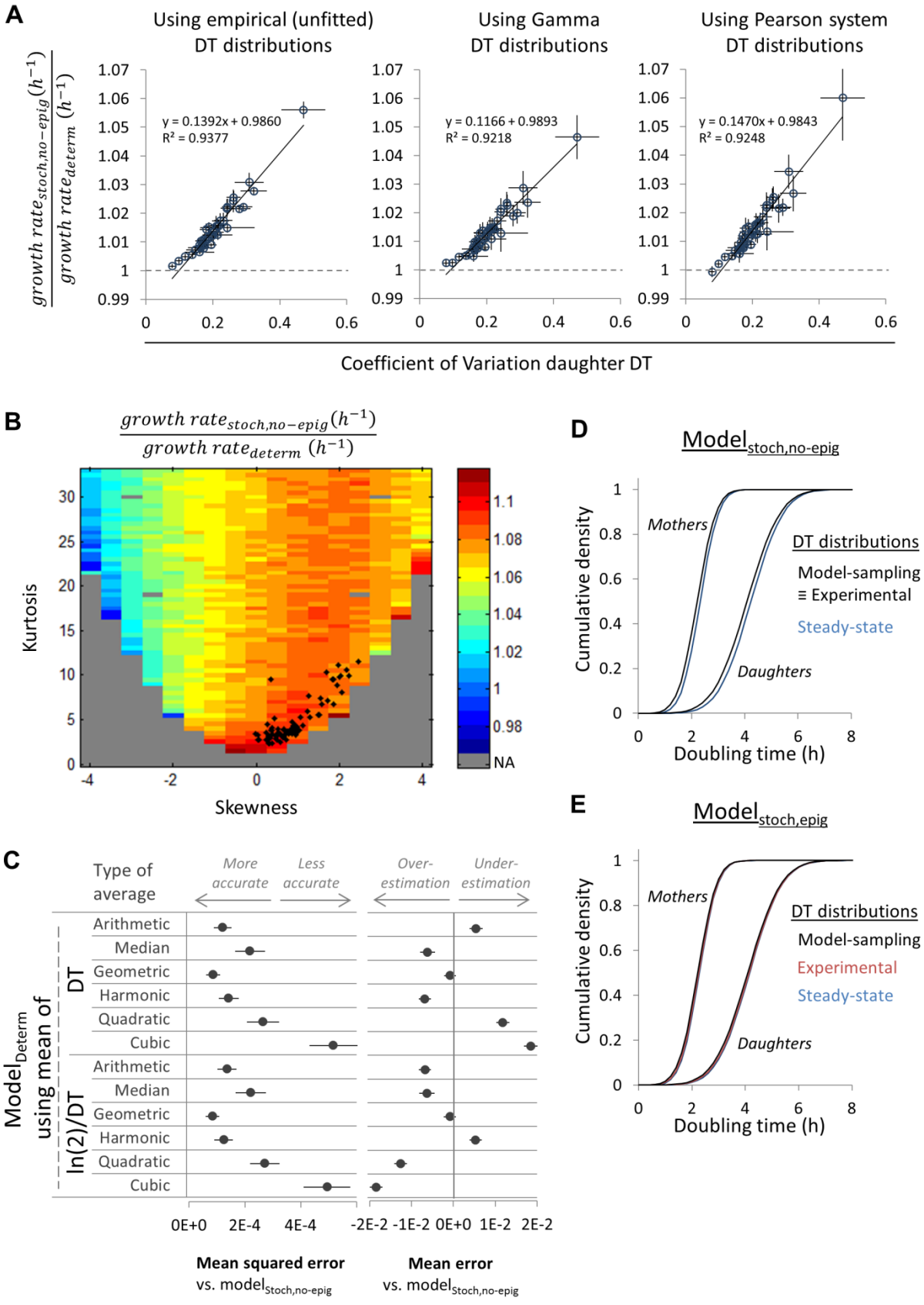

A
Using empirical (unfitted) DT distributions
Using Gamma DT distributions
Using Pearson system DT distributions
y = 0.1392x + 0.9860
R² = 0.9377
y = 0.1166 + 0.9893
R² = 0.9218
y = 0.1470x + 0.9843
R² = 0.9248
Coefficient of Variation daughter DT
B
Kurtosis
Skewness
C
Type of average
More accurate
Less accurate
Over-estimation
Under-estimation
Arithmetic
Median
Geometric
Harmonic
Quadratic
Cubic
Model_Determ using mean of
DT
ln(2)/DT
Mean squared error vs. model_Stoch,no-epig
Mean error vs. model_Stoch,no-epig
D
Model_stoch,no-epig
Mothers
Daughters
DT distributions
Model-sampling ≡ Experimental
Steady-state
Cumulative density
Doubling time (h)
E
Model_stoch,epig
Mothers
Daughters
DT distributions
Model-sampling
Experimental
Steady-state
Cumulative density
Doubling time (h)

Figure S4. Related to Figure 4. The validity of the use of normal distributions in the model and the legitimacy of our approach to measure its parameters.

(A) DT variance increases population growth rate in simulations of growth driven by (left) the empirical DT distribution, (middle) Gamma DT distributions and (right) Pearson system DT distributions. Similar to Figure 4A, the ratio of predicted population growth rates of $model_{stoch,no\text{-}epig}$ over the deterministic model is plotted against the daughter DT CoV for all experiments in the dataset. In contrast to Figure 4A, simulations were performed using sampling distributions that differ from normal distributions. Error bars represent bootstrapped SDs (Supplemental Experimental Procedures).

(B) Skewness and kurtosis in the model-sampling distribution can affect the population growth rate predictions. The effect of skewness and kurtosis was modeled using Pearson system distributions. This is a family of probability distributions that allow the construction of distributions given a combination of the mean, variance, skewness and kurtosis. All simulation were performed using mother DT mean = daughter DT mean = 4, and mother DT SD = daughter DT SD = 2. The black diamonds represent the skewness and kurtosis values that were observed in our dataset (Dataset S1). Normal distributions have skewness = 0 and kurtosis =3.

(C) Population growth rates were simulated for all experiments in the dataset with the deterministic model using a range of different averages (the median and the arithmetic, geometric, harmonic, quadratic and cubic mean). These averages were taken directly on the DT measurements, or after converting them to 'growth rate'-values [using growth rate = ln(2)/DT] followed by conversion of the average back to a DT-value. We assessed which mean would best predict population growth, by calculating the mean squared error (left) and mean error (right) compared to the stochastic model. Importantly, the arithmetic mean of the 'growth rate'-values of the sampling distribution overestimates the population growth rate, since the steady-state distribution is enriched with slowly-growing cells (see Results and Supplemental Experimental Procedures). This analysis also indicates that the geometric mean of the single-cell DTs or ln(2)/DT's are good predictors of population growth rate under a deterministic model.

(D) and (E) The legitimacy of using our experimentally measured DT parameters as the model's parameters. Three different DT distributions can be described in the stochastic model: the (model^{-})sampling, the steady-state and the experimental distribution (see Supplemental Experimental Procedures for definitions). All simulations were performed using the experimentally measured parameters from BY/S288c growing in palatinose 2%. (D) The three cumulative DT distributions under our model without epigenetics ($model_{stoch,no\text{-}epig}$). The experimental and the model-sampling distribution are equal by definition, since the model explicitly assumes that all newly-assigned DTs are independent of the previously assigned DTs. The steady-state distribution is enriched for cells that have been assigned a long DT. (E) The three cumulative DT distributions under the model including epigenetics ($model_{stoch,epig}$). When taking into account epigenetic inheritance, the steady-state and model-sampling distributions become more similar to each other. In contrast to the model without epigenetics, there is a slight difference between the model-sampling and the experimental distributions.

Figure S5

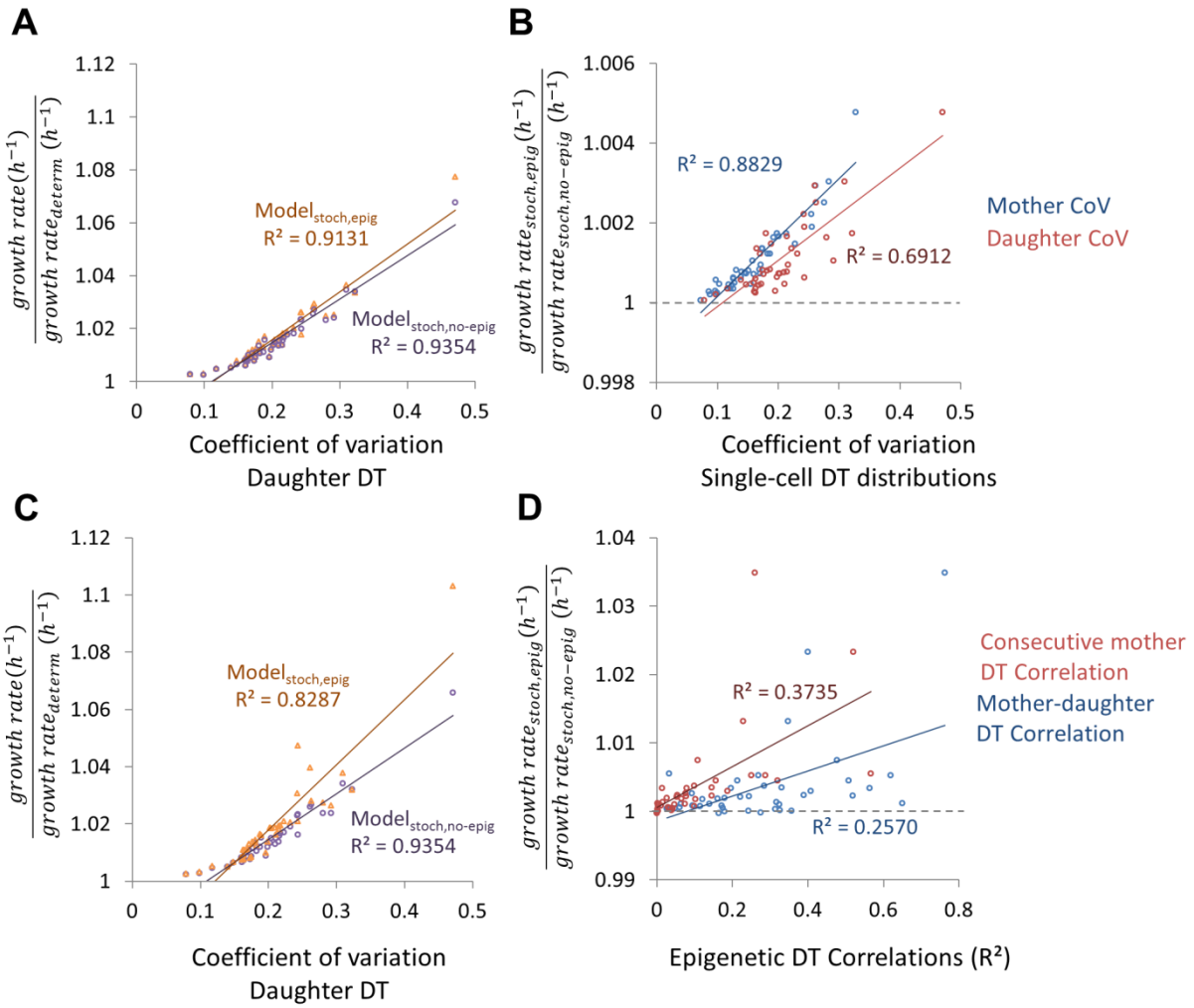


Figure S5. Related to Figure 5. Epigenetic inheritance and single-cell DT variability both increase population growth rate predictions when epigenetic effects are modeled.

(A-B) Predicted population growth rates using the model with average epigenetic DT inheritance across all experiments. The R^2 of the DT correlations used in these simulations were average correlations across all experiments (Figure 2D and Figure S1). (A) Fractional growth rate compared to the deterministic model is plotted against variability in daughter DT distributions, for 1) the model without epigenetics (purple circles; as reference) and 2) the model with epigenetics (orange triangles). (B) The fractional growth rate increase when epigenetic effects are considered is plotted against variability in mother (blue) and daughter (red) DT distributions.

(C-D) Predicted population growth rates using the model with individual (per experiment) epigenetic DT inheritance. The R^2 of the DT correlations used in these simulations were calculated for each experiment individually (Figure 5A). (C) Fractional growth rate compared to the deterministic model is plotted against variability in daughter DT distributions, for 1) the model without epigenetics (purple circles; as reference) and 2) the model with epigenetics (orange triangles). (D) The fractional growth rate increase when epigenetics effects are considered is plotted against 1) the correlation between consecutive mother DTs (red) and 2) the correlation between a mother and its most recently born daughter (blue).

Figure S6

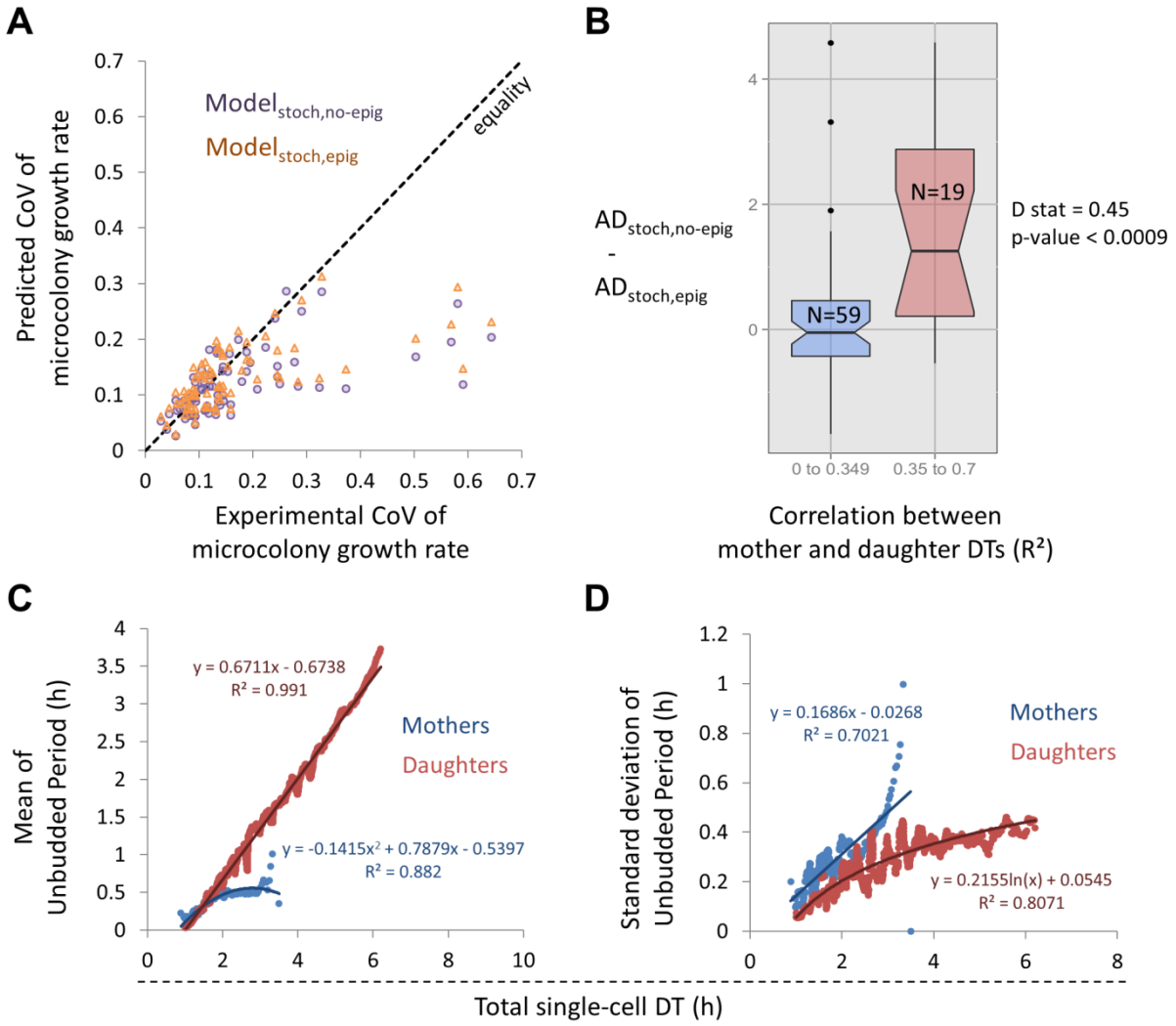


Figure S6. Related to Figure 3. Modeling microcolony growth rate variability: incorporating cell-cycle dependent single-cell growth into the model.

(A-B) The stochastic model accurately predicts microcolony growth rate variability at small population sizes. All the model's predictions were done using the individual (per-experiment) epigenetic DT inheritance. (A) Predicted vs. experimentally determined microcolony growth rate variability (as coefficient of variation; CoV) using the model (with or without epigenetics). (B) When epigenetic DT inheritance is high, the model's predictions are more accurate when epigenetic effects are modeled. Shown are boxplots for the experimental conditions with the lowest and highest mother-daughter DT correlations. Each data point is the Anderson-Darling (AD) statistic of a comparison between the empirically observed microcolony growth rate distribution compared to the purely stochastic model's prediction minus the same test performed against the stochastic-epigenetic model's predictions (all growth rates normalized to the median growth rate within the distribution). Note that positive values of $AD_{stoch,no\text{-}epig}$-$AD_{stoch,epig}$ indicate a better fit when epigenetic effects are considered in the model.

(C-D) To model microcolony growth rate variability in (A), the model was extended to also take into account cell-cycle dependent single-cell growth (see Supplemental Experimental Procedures). The following empirical relationships were used to stochastically model the length of the unbudded and budded period in each single cell. (C) The mean and (D) SD of the unbudded period is plotted as function of the total DT for mother (blue) and daughter (red) cells. Mean and SD were calculated using a sliding window of 50 observations, after sorting the data points from smallest to largest DT. The number of observations for mother and daughter cells was respectively 693 and 2547.

Supplemental Experimental Procedures

1 Yeast strains and growth media

1.1 Yeast strains and growth media used (extended)

The specific strain and media combinations used for the 41 experiments in our dataset can be found in Dataset S1. Several experiments were performed using BY/S288c-derived strains [S1]. These include the strains AN148 and AN296, both of which were derived from a specially engineered maltose-prototrophic S288c strain, in which a functional MAL regulator allele was introduced in place of MAL13 on chromosome VII starting from the strain BY4741 [S2]. Furthermore, in these strains, the high petite frequency that occurs after extended growth on glucose, which is characteristic to S288c, was reduced by rescuing a frameshift mutation in SAL1 [S3]. And finally, a constitutive YeCitrine marker was introduced into AN296; and fusion constructs of MAL11-YeCitrine (encoding a maltose transporter) and MAL12- mCherry (encoding a maltase) were introduced into AN148. Another strain used in our experiments, AN146, was constructed by crossing a low-petite frequency S288c-strain (made low-petite frequency as described above) with a maltose-prototrophic (made maltose-prototrophic as described above) derivate strain of UCC8363 [S3].

The strains AN148 and AN296 were used as progenitor strains in separate experimental evolution experiments in [S4], resulting in a series of isolates with altered growth characteristics Some of these isolates were included in our experimental dataset (Dataset S1). The progenitor AN296 gave rise to an isolate called AN296_1 (evolved isolate 1 in Figure 3 of [S4]). The progenitor AN148 gave rise to the isolate AN148_pop6r6 (evolved isolate 5 in Figure 3 of [S4]) and a series of isolates coming from 12 independently evolving populations. The latter were termed AN148_x.y with x representing the population and y representing the first, second or third isolate from this population [S4].

1.2 Strain construction details

For the flow cytometry analysis shown in Figure 6B, the following strains were constructed. The wild-type (WT) strain BY4741 [S1], was used to construct the fluorescently tagged strain KP30, which contains fusion constructs of *MAL11*-mCherry and *IMA1*-yECitrine. Similarly, the WT strain was used to construct KP4, a strain which contains the same fusion constructs, but also an extra copy of *MAL11* expressed from a GPD promoter that was introduced in a neutral locus in the genome (chromosome XIII in between *ATG16* and *YMR158C-A*; coordinates: 574155-573795). For the experiments shown in Figure 6C, the WT strain was compared with a strain (KP198) only bearing the *MAL11* overexpression, without the fluorescent protein fusions.

Strain list for strains used in Figure 6. Primers and plasmids used for construction of the strains are listed.

Strain Number	Parental Strain	Genotype difference with parent	Forward Primer for Amplification	Reverse Primer for Amplification	Template	Check Primer 1	Check Primer 2	Selection Marker
BX1	BY4741	IMA1-yECitrine	KP6	KP7	pKT139 (Euroscarf)	KP1	KP8	HIS5
KP30	BX1	MAL11-mCherry	AN14	AN15	pSR101 [S2]	KV1203	KV69	URA3
KP198	BY4741	pGPD-MAL11	KP85	KP86	gDNA AN104 [S4]	KP17	KV69	KAN
BX20	KP198	IMA1-yECitrine	KP6	KP7	pKT139 (Euroscarf)	KP1	KP8	HIS5
KP4	BX20	MAL11-mCherry	AN14	AN15	pSR101 [S2]	KV1203	KV69	URA3

Sequences of primers used for strains used in Figure 6.

Primer Name	Sequence
KP6	gaggtagatgcctcttccagaacattgaagccatgggaaggaagaatatatatcagcgaaATCGGTGACGGTGCTGGTT

KP7	ctaaaacatcacttttgtagggtttcttcgcacattatcattattattctttgagaatacGAATTCGAGCTCGTTTAAACTGG
KP8	attcatcgttcaagttagagtttgg
KP1	taagtagcatcaccttcaccttc
AN14	gagagtatcagtcagtcctcaagcataaaacagcgagaattaaatgcagctgataaatgtATCGGTGACGGTGCTGGTTT
AN15	tttgggagcagtcaaagggattccttatttcttccaaaaaaaaaaaaacaacccttttacgcATAGGCCACTAGTGGATCTG
KV1203	ggcttgagggacatacagaga
KV69	gcacgtcaagactgtcaagg
KP85	ccactcacgccaatatactaaccaattgaaatgctcataaccagaatgaacaggccgcatgcattc CGTACGCTGCAGGTCGAC
KP86	gtatcttgagagaaactcaatagtaataactggatgcataaagaaaatttcatatgaatcactac GTTGTTCGAATCAATATCCAGGCAC
KP17	tcagcgagaggatcagcaac

2 Time-lapse microscopy: movie acquisition and analysis

2.1 Acquisition of time-lapse movies (extended)

The acquisition of the movies was done automatically using the NIS-Elements software (version 3.2) in combination with a Nikon Eclipse Ti microscope equipped with a DL-604M-#VP camera (Andor™ technology) placed in a temperature-controlled incubator (30°C). All movies were captured using a 60x 1.4 NA oil immersion lens, and proper focus of the images was controlled using the microscope's automated Z-plane focusing. The time interval with which pictures were taken was based on the population growth rate of the examined strains, and ranged from 3 (very fast growth) to 20 (very slow growth) minutes.

2.2 Analysis of single-cell DTs (extended)

Within one microcolony starting from a single cell, we were generally able to measure two mother DTs and two daughter DTs, before the cells within the microcolony became too dense to be scored (Figure 1A). All DTs were measured by scoring the time period between cell divisions (a cell division causes the bud to become physically separated from its mother leading to a sudden shift in its position). In addition, we also scored unbudded and budded periods for daughter cells, by scoring the time of first bud appearance. This allows a division of the DT into an unbudded and a budded period (Figure 1A). The extent to which human error introduces noise is limited, as separate investigators independently replicated results of preliminary analyses.

Per experiment, we have calculated mean, SD, CoV, skewness and kurtosis for mother and daughter DT distributions using a bootstrapping (resampling) approach. Raw DT data were resampled 1000 times with replacement using an N that equals the total number of DT observations in each specific experiment. For each resampled population of DTs, a Normal distribution was fitted in MATLAB using the fitdist() function, and mean, SD and CoV were extracted. The skewness and kurtosis were calculated in MATLAB using the functions skewness() and kurtosis(), in both cases correcting for sample bias by setting flag=0. For each parameter, the 'bootstrapped mean' and 'bootstrapped SD' were calculated. These 'bootstrapped SD'-values were used to generate error bars in Figures 2A,C and Figures 4A and S4A.

For 21 out of 41 experiments, we have scored the unbudded period of mother cells (not shown in Figure 1A) in a similar fashion, by scoring the bud appearance in the 1st and 2nd daughter cell after its first cell division. At this point, these cells have already undergone one cell division, and have thus effectively become a mother cell.

All the above raw data and statistics calculated can be found in Dataset S1 for all 41 strain/condition combinations used in our dataset, and the results shown exclusively in Figure 6C.

2.3 Fitting distributions to DT data and testing goodness-of-fit

To develop a single-cell-based model for population-level growth that is capable of incorporating the effect of single-cell heterogeneity, we examined how to best mathematically approximate the measured DT distributions.

First, different probability distributions were fitted to the DT data. All mother and daughter DT measurements were grouped per experiment, and Normal, Gamma, Weibull, Log-logistic, Log-normal and Exponential distributions were fitted to them using a bootstrapping (resampling) approach. In short, the DT data were

resampled 1000 times with replacement at full population size and distributions were fit using the fitdist() function in MATLAB (Dataset S2A).

Next, the goodness-of-fit was examined for the different fitted distributions (Normal, Gamma, Weibull, Log-logistic, Log-normal and Exponential). For each bootstrapped set of DTs, Anderson-Darling Tests for goodness-of-fit of the different fitted distributions (see above) were performed using the function adtest() in MATLAB. As the DT data was resampled 1000 times per experiment, this led to 1000 p-values per experiment per type of distribution. To assess whether a certain distribution gave a good fit with our experimental data, an average p-value was calculated as the mean of these 1000 'bootstrapped p-values' (Dataset S2B).

We found that for the vast majority of strain/condition combinations, a Normal distribution provides a good fit (Bonferroni corrected p-value>0.05 for 39 of 41 mother DT distributions, and 41 of 41 daughter DT distributions). Only the Gamma distribution provides a marginally better fit (Bonferroni corrected p-value>0.05 for 40 of 41 mother DT distributions, and 41 of 41 daughter DT distributions; Dataset S2B).

Since the Normal distribution is a probability distribution with very intuitive parameters, we decided to model population growth using the stochastic models under the assumption that the DT distributions are well described using Normal distributions. However, since Gamma distributions fit our experimental data marginally better than Normal distributions, and some of the DT distributions show a clear positive skew (Figure S4B), we investigated how deviations for normality influenced the population growth rate predictions (see Results section and Figure S4A-B).

3 Empirically measuring population growth rates

3.1 Measuring population growth rates on solid media

Populations growth rates on solid media were measured as the increase of cell count (Figure 3B) or microcolony area (Figure S3B) over time. Per time-lapse movie, microcolonies were analyzed by measuring the cell count (or area) at three evenly spaced time-points throughout the movie, in such a way that the total area of the microcolonies increased between 2-18 fold from the first timepoint to the last (Dataset S3,5). Next, the cell count (or area) of all the microcolonies at each timepoint was summed and the microcolony growth rate was calculated as the slope of a linear model fitted to the ln-transformed summed cell count (or area) plotted against time (Dataset S3). We measured the microcolony area by fitting regions of interests around the microcolonies using the NIS-Elements software. The number of cells (each visible cell shape, including buds) within the microcolonies was scored manually.

3.2 Measuring population growth rates in liquid media

Growth rates in liquid culture (Figure 3C) were measured using colony counting. Strains were inoculated into glass tubes containing 3mL of the appropriate medium and grown overnight to turbidity in a spinning wheel at 30°C. The next day, a dilution series was made from the overnight cultures into a microplate and incubated on a plate shaker at 30°C. After 6 hours of adaptation to diluted growth conditons, dilutions within the OD_{600} range of 0.1 to 0.3 were taken and inoculated into a falcon tube containing 3mL of medium so that they would still be dilute the next day. Colony forming units were used to determine population sizes. These colony forming units were obtained immediately after dilution into the falcon, and finally after 16-24 hours of growth. Growth rates were calculated as $\frac{(\ln\langle pop.size_{final}/pop.size_{initial}\rangle)}{(t_{final}-t_{initial})}$.

4 Predicting population growth rates

4.1 Predicting population growth rates using the stochastic model (extended)

The model keeps track of mother and daughter cells in two matrices, with each row containing one cell and each column keeping track of a single-cell parameter. At birth or after the completion of a bud, each cell is assigned a new DT which is dependent on the experimentally measured DT variance and DT correlations within lineages (see Results, Figure 3A and below). The first parameter that the model keeps track of is the time tau remaining for completion of the cell's bud. The second parameter is the last DT that was assigned to that cell. The model iterates with respect to time (in steps of 0.001h), subtracting the elapsed time t from each cell's tau value. When a cell's tau equals zero, a new daughter is born, which then enters into the daughter compartment with a tau equaling the newly assigned DT. After producing a new daughter, daughter cells enter the mother matrix whereas mother cells simply remain within this compartment. Simultaneously, these cells are assigned a tau equaling the newly assigned DT.

In order to simulate exponential population growth, we grow the cells (in silico) in the following way. First, a population of 10,000 cells is grown starting from 1 mother and 1 daughter cell. Next, we randomly pick 100 cells from this population and grow them to a size of 100,000. In this way, we reach a steady-state at the end of the simulation, independent of the initial starting population. We then calculate instantaneous growth rates across the whole growth curve by calculating the growth rates over windows of 10 min. Finally, the population growth rate is calculated by averaging the instantaneous growth rates measurements acquired during the last hour of growth before the population reaches 100,000. The script for the stochastic model(s) is written in the MATLAB language (MATLAB version 8.1.0).

In some cases, the simulations have difficulties converging to a stable steady-state when starting from an initial population containing 1 mother and 1 daughter cell. This occurs when the DT variability is not large enough to 'desynchronize' growth within the population, leading to staggered or synchronized growth patterns. This phenomenon occurs when the simulations are performed using the empirical DT distributions (Figure S4A), or when the CoV in both the mother and daughter DT distributions is lower than 10% (Figure 4B). Therefore, in these specific cases, we have started the simulations using initial population consisting of 100 mother and 100 daughter cells, which are given initial tau's ranging from 0 until the mean DT in equal steps of (mean DT)/100. This approach allows the simulations to reach a steady-state growth rate, when the DT variability is too low.

All growth rate predictions of the stochastic model were obtained using a bootstrapping (resampling) approach. The raw DT measurements were resampled with replacement at their full size and Normal distributions were fitted to them using fitdist() in MATLAB. The means and SDs of these distributions were then used as parameters for the simulations, and growth rates were obtained as explained above (as the average of 3 repeated simulations). For each experiment, resampling was repeated 100 times, leading to means and SDs for the simulated growth rates. These 'bootstrapped SD'-values were used to generate error bars in Figures 3B-C, 4A, S3B and S4A.

The simulations using other distributions than the Normal distribution (Figure S4A) were obtained using the same bootstrapping approach. However, when using Gamma distributions, the fitdist() function was used to extract the shape and scale parameters. When using Pearson system distributions, the first four statistical moments were calculated using the mean(), var(), skewness() and kurtosis() functions, with corrections for sample bias by setting flag=0 in the case of skewness and kurtosis. Finally, when using the empirical DT distributions, we did not fit statistical distributions, but sampled directly from these distributions during the simulations.

4.2 Predicting population growth rates using the deterministic model (extended)

Population growth rate predictions using the deterministic model were calculated in parallel using the same bootstrapping approach as used for the stochastic model (see above), using the same resampled DT distributions.

4.3 Simulations of selection for mutants with altered DT characteristics using the stochastic model.

We have used the stochastic model (without epigenetics) to test if cells that acquire a mutation leading to a higher fitness (predicted growth rate) can rise in frequency in the population due to selection. Specifically, we tested if the efficacy of selection is dependent on whether the fitness increase occurs through changes in the DT mean or DT noise. Three hypothetical strains were used in these simulations: 1) The wild-type strain (WT), or the background population in which spontaneous mutants can arise, 2) The "variable-mutant" genotype (VarMut), which has increased fitness compared to the WT, due to increased DT noise, while keeping the mean DTs unchanged, and 3) The "homogeneous-mutant" genotype (HomMut), which has similarly increased fitness compared to the WT, due to shorter mean DTs, while keeping the noise (CoV) constant.

The DT parameters used for the WT are mother DT ~ $N(2,0.1)$ and daughter DT ~ $N(4,0.1)$. The DT parameters used for the VarMut are mother DT ~ $N(2,0.3)$ and daughter DT ~ $N(4,0.3)$. The DT parameters used for the HomMut are mother DT ~ $N(1.937,0.1)$ and daughter DT ~ $N(3.874,0.1)$. The predicted population growth rates of the three strains are respectively 0.241±0.001, 0.249±0.001 and 0.249±0.001 for the WT, VarMut and HomMut. Note that the single-cell parameters of the VarMut and the HomMut were picked in such a way that they exhibit the same fitness increase compared to the WT (fitness[VarMut]=fitness[HomMut]=1.032).

Using these mutants, we examined the probability of fixation and extinction of spontaneously arising mutant cells in a population of wild-type cells. Specifically, we simulated growth of a population of WT cells seeded

with 1 mutant cell. Initially, we started with exponentially growing populations of 1000 WT cells seeded with 1 mutant cell. Each time the total population reached 2000 cells, it was diluted again by random sampling to 1000 cells. The simulations were stopped either when the mutant genotype reached 10% fixation, when the mutant genotype became lost from the population, or after 1000h of growth when none of the two previous events occurred. All strains from were competed against the WT for 10 000 repeated simulations (WT vs. WT, VarMut vs. WT, HomMut vs. WT). Under these conditions, we find that the "variable-mutant" genotype and the "homogeneous-mutant" genotype have very similar chances of becoming fixated in the population. More specifically, the VarMut and the HomMut respectively have a chance of 7.53% and 7.94% to become fixated in the population, and a chance of 92.5% and 92.1% to become extinct. None of the simulations were undecided after 1000h of growth. As a control, we also seeded WT cells to see what is the chance to become fixated in the population by chance (drift). For WT cells, 0.89% of the simulations lead to fixation, 98.9% leads to extinction and 0.23% is undecided after 1000h of growth.

4.4 Modeling growth rate variability at small population sizes

For the predictions of microcolony growth rate variability, an extended version of the growth model was used, which also takes into account growth of single cells as they progress through the cell cycle (section 8). Growth simulations were carried out in a similar fashion as described in section 4.1. First, a steady-state population of 100,000 cells was grown starting from 1 mother and 1 daughter cell. Next, in each simulation 1000 random cells were sampled individually from this population and allowed to grow for 1.5-12h, mimicking the lengths of time of growth for the empirical microcolony area growth rate measurements (Dataset S5). Then, total growth rates were calculated using Supplemental equations 23 and 24.

5 Statistical analyses

5.1 Population growth rates measured by tracking microcolony areas consistently underestimate population growth rates

We have measured population growth rates for the strain/medium combinations used in our dataset (Dataset S1). These growth rates were measured both on solid medium (Figure 3B and S3B-C) and in liquid medium (Figure 3C). Moreover, the growth rates on solid medium were measured by analyzing the time-lapse movies in two different ways: either by tracking the increase in microcolony area, or by tracking the increase in cell count within the microcolonies (section 3). While we find that these two measures are highly correlated ($R^2 = 0.9709$; Dataset S3), we also find a consistent discrepancy between them (Figure S3B). When using microcolony area growth rates, the resulting population growth rate is on average 13.5% lower (ranging between 30.91% lower and 5.4% higher) compared to the microcolony cell count growth rates. Indeed, when we compare the increase in area and cell count within the exact same microcolonies, we find that the area increases more slowly than the cell count (Figure S3C).

We reason that growth rates measured using microcolony area increases consistently give lower growth rate estimates, since adhesive forces on agar could restrict the outwards expansion of these microcolonies. Cells growing in microcolonies would then be pressed together, leading to overlapping of the cell's surfaces. This would lead to a slower increase of the microcolony areas, and thus an underestimation of population growth rates.

In Figure S3A, we compared the predictions of our model (run under the different assumptions shown in Figure 3A) with the population growth rate measurements. While we find that the model's predictions show a strong correlation with both the area and cell count population growth rates, the model more accurately predicts the cell count growth rates (Figure S3A,B). Similarly, the model also accurately predicts the growth rate measured in liquid (Figure 3C and S3A). Together, this again indicates that population growth rates measured by tracking microcolony areas consistently underestimate population growth.

5.2 Statistical analysis of the ability to predict microcolony growth rate variability

In Figure S6, modeled microcolony growth rate variability was compared to empirical variability (microcolony area growth rates). In total, N=78 empirical distributions were compared with the modeled distributions from simulations using $model_{stoch}$ and $model_{stoch,epig}$. These empirical distributions correspond to 40 unique strain/condition combinations, with maximum 2 population sizes per strain/condition (1-13 cells, median 3.3 cells; see Dataset S5)

The modeled microcolony growth rate distributions were compared to the empirical distributions using the Anderson-Darling (AD) Test. In order to allow comparison of these distributions, all microcolony area growth rates were normalized to within-distribution median values. This is necessary since microcolony area growth rates are consistently underestimating actual growth rates (Figure S3). Of course, in this way we are only testing if microcolony variability is accurately predicted by the models. However, this approach is justified since we have already shown that the model can accurately predict average population growth rates (Figure 3 and S3). Per AD test, there were between 19 and 62 (median 32) empirically measured microcolony area growth rates compared to a set of 1000 simulated microcolony growth rates. Each test results returns a p-value and an AD-statistic which reflects the extent to which the two distributions are different from one another (Dataset S5). To test which model was more accurate, we used a linear regression model in R using the formula lm(difference ~ MDR2) where 'MDR2 equals the vector of correlations for each strain and condition's mother-daughter DT correlation and 'difference' is the vector of values equal to the AD-statistics of the empirical vs. $model_{stoch,no\text{-}epig}$ tests minus the AD statistics of the empirical vs. $model_{stoch,epig}$ tests. Note that positive value of this 'difference' indicate a better fit of the $model_{stoch,epig}$ to the empirical data compared to the $model_{stoch}$. We tested other models that included population size and mixed effects of strain (S288c binary variable) and repeated measures and found similar significant results for MDR2. In Figure S6B, the test results were binned according to the mother-daughter DT correlation ($0<R^2<0.349$ and $0.35<R^2<0.37$), and 'difference' was plotted per bin using boxplots. A KS-test of the two distributions in these bins indicates that the these two 'difference' distributions are significantly different from each other ($D = 0.45$, $p< 0.0009$).

5.3 Statistical analysis of the relationship between single-cell and population-level statistics and simulations

The data in Dataset S5 containing bootstrapped means for single-cell and population-level statistics and simulations were used to explain how life-history traits (DT-related statistics) could predict population-level behavior. Linear regression using R's lm() function was used for all modeling. One key model included which variables best predicted population-level growth rates. Over all statistical moments of mother and daughter DTs, variation in mean DT was sufficient to explain most measured population-level variation. Another key model was to look at the fold increase in predicted growth rate of stochastic models over determinstic models. This fold-increse number was best predicted in a linear model by a model with mother and daughter COVs as explanatory variables. Mother- and daughter-cell COV has a slight dependency on mean variation in DTs, and therefore this analysis was repeated by using residuals of in lm(COV~mean) model in a new linear model to predict the predicted fold-increase in growth rate.

6 Flow cytometry to measure gene expression noise

For each trace shown in Figure 6B, a total of 50,000 single-cell events were acquired by a BD Biosciences Influx flow cytometer. For mCherry signal detection, a 561 nm laser coupled to a 610/20 nm detector was used. For yECitrine signal detection, a 488 nm laser coupled to a 530/40 nm detector was used. R's flowCore package was used to first filter out ~20-30% of events using a filter (curv2Filt) that selected the highest-density regions in side- and forward-scatter dimensions. Each trace shown in Figure 6B shows only these filtered events. mCherry and yECitrine events for each filtered sample were normalized to by FSC-median-normalized FSC signal, that is, for each i-th event, signal_normalized_i = signal_raw_i / (FSC_raw_i/median_FSC_signal) [S5]. These FSC-normalized values were stored as binned fluorescent measurements and summary statistics.

7 Constructing a stochastic model which takes into account epigenetic DT inheritance ($model_{stoch,epig}$)

In the following sections, we will first show correlations between the DTs of closely related cells within lineages (section 7.1 and 7.2). This analysis will be done across the full dataset, combining all single-cell measurements in these correlations (Dataset S1). Next, we will construct a stochastic model that takes into account (some of) these correlations. In contrast to a purely stochastic model ($model_{stoch,no\text{-}epig}$), the choice of new DTs will be dependent on previously assigned DTs of closely related cells. This should then lead to correlations between the DTs of these cells. First, we will derive the necessary mathematical equations to do this (section 7.3). Then, this set of equations is built into our stochastic model (section 7.4). And finally, population growth rates are predicted using this complete model which incorporates experimentally observed epigenetic DT inheritance (section 7.5).

7.1 Correlations between DTs within cell lineages reveal epigenetic DT inheritance

During analysis of the time-lapse growth records, we scored DTs of mother and daughter cells within small microcolonies. Within each microcolony, we have tracked mother cells for two consecutive DTs, and scored the DTs of two daughter cells that it produced (Figure 1A and Figure S1A). Using these relationships, we investigated if DTs of related cells within lineages are correlated (epigenetic DT inheritance, see Result section).

In an initial approach, we checked for correlations between the measured DTs across the whole dataset (n=41 experiments; Dataset S1). In this approach, all single-cell measurements were combined and no corrections were made for the differences in number of DTs scored between the different experiments. This leads to some sort of average degree of epigenetic DT inheritance across the whole dataset. However, we do note here that there are large differences in epigenetic DT inheritance between different experiments (see Figure 5A).

In order to be able to compare across the whole dataset, we first transformed all measured DTs to z-scores (equations 1 and 2). For each DT, z-scores were calculated relative to the fitted normal distributions that were determined for mother and daughter cells within each specific experiment (Dataset S2). In this way, an observation with a z-score higher than 0 always shows a DT which is longer than the average DT within the mother or daughter fraction, while z-scores lower than 0 indicate DTs shorter than the average.

$$mDT_{i,zscore} = \frac{(mDT_i - mDT_{mean})}{mDT_{stdev}}$$

$$mDT_i = \ i^{th}\ mother\ DT\ (i \in [1,2,3,..,n]\ with\ n = total\ observations)$$

$$mDT_{i,zscore} = z\text{-}score\ of\ the\ i^{th}\ mother\ DT$$

$$mDT_{mean} = \ mean\ of\ fitted\ mother\ DT\ distribution$$

$$mDT_{stdev} = \ standard\ deviation\ of\ fitted\ mother\ DT\ distribution \quad \textbf{(1)}$$

$$dDT_{i,zscore} = \frac{(dDT_i - dDT_{mean})}{dDT_{stdev}}$$

$$dDT_i = \ i^{th}\ daughter\ DT\ (i \in [1,2,3,..,n]\ with\ n = total\ observations)$$

$$dDT_{i,zscore} = z\text{-}score\ of\ the\ i^{th}\ daughter\ DT$$

$$dDT_{mean} = \ mean\ of\ fitted\ daughter\ DT\ distribution$$

$$dDT_{stdev} = \ standard\ deviation\ of\ fitted\ daughter\ DT\ distribution \quad \textbf{(2)}$$

We then correlated these z-scores for all possible genealogical relationships within our lineages (Figures S1B-G). We find the strongest (average) correlation for $dDT_{1,zscore}$ vs. $mDT_{1,zscore}$ (R^2=0.2627) and $dDT_{2,zscore}$ vs. $mDT_{2,zscore}$ (R^2=0.1497). Note that both these correlations represent the DTs of a mother and her most recently born daughter (Figure S1A). The remaining correlations are shown in Figure S1D-G, for which the R^2-values range from 0.0030 to 0.0652. For consecutive mother DTs, we find an (average) correlation of R^2=0.0153 (Figure S1D).

7.2 Analysis of the DT correlation between a mother and her most recently born daughter

We decided to explore the DT correlations shown in Figures S1B-C more thoroughly, since these gave the highest R^2-values. Note that both panels in this figure show the DT correlation between a mother and her most recently born daughter. Since these data essentially show the same genealogical relationship, we decided to combine them for further analysis. This combined correlation ($dDT_{1/2,zscore}$ vs. $mDT_{1/2,zscore}$) is shown in Figure 2d, and has an R^2-value of 0.2517. In order to get a clear view on the density of observations in this scatter plot, the histograms of the $mDT_{1/2}$ and $dDT_{1/2}$ z-scores are shown in Figure S1H-I.

To further characterize the relationship shown in Figure 2d, we have binned the data on the x-axis ($mDT_{1/2,zscore}$) in 19 equally sized bins, and fitted normal distributions to the resulting y-axis values in each bin

($dDT_{1/2,zscore}$). In Figure S1J, the means and SDs of these fitted distributions were plotted against the mid of each bin.

The most straightforward approach to build this DT correlation into our stochastic model, would be to sample all daughter DTs from normal distributions for which the parameters depend on the mother DTs as shown in Figure S1J. As in the purely stochastic model, the mother DTs would still be sampled randomly from the experimentally determined mother DT distributions. From the histograms shown in Figure S1H-I, it becomes clear that the average DT distribution has a small right-handed tail (positive skewness). This asymmetry is reflected in the higher observed means and SDs of the fitted $dDT_{1/2,zscore}$ distributions for positive $mDT_{1/2,zscore}$ values (see Figure S1J). However, in our stochastic growth model, we approximated the DT distributions with normal distributions (see Result section). Therefore, we cannot simply use the relationships shown in Figure S1J to incorporate this DT correlation into our purely stochastic growth model. Instead, in what follows we will derive these relationships for the idealized situation in which mother and daughter DT are normally distributed.

7.3 A mathematical framework to model two correlated normal distributions

As explained above, we cannot directly use the relationships shown in Figure S1J, since these apply to the observed DT_{zscore} distributions which have a slight right-hand tail (positive skewness). Moreover, we would like to be able to build in different strengths of correlation ($0<R^2<1$) between DTs. Therefore, in the following sections, we will work out a general mathematical framework which allows us to model specific strengths of correlations (R^2) between two normal distributions of DT_{zscore}'s which can have any given genealogical relationship. First, we will explain the general approach to achieve this, without elaborating on the mathematics behind it (section 7.3.1). Next, we will derive the necessary mathematical expressions, first using an exact mathematical approach (section 7.3.2), and second using a trial-and-error simulation approach (section 7.3.3).

7.3.1 General concept

During *in silico* growth in our model, new DTs need to be assigned every time a cell is born or completes its doubling. Let's imagine a population of cells which has DT_{zscore}-values following a standard normal distribution. For each of these cells, a new DT_{zscore}-value has to be picked when their DT has finished. When this is done for all cells, we should again end up with a standard normal distribution of DT_{zscore}-values. This is a necessary prerequisite for incorporation of epigenetic effects into our purely stochastic model, and the only condition that we impose on the set of equations that will be used to determine the new DTs. In what follows, these initial and final DT_{zscore} distribution will be referred to as respectively the input and output distribution. The function which is then used to determine the new DT_{zscore} for each cell individually will be called the transfer function.

In theory, we can now distinguish three situations, which are represented in Figure S2A. The first situation (left panel) is simulated in the purely stochastic model. Here, the new DT_{zscore}'s are sampled independently of the previous DTs. The transfer function (used to determine a new DT_{zscore} for each cell individually) in this case is simply the standard normal distribution N(0,1). There will be no correlation between the input and output distributions, leading to a model which does not incorporate epigenetic DT inheritance. When this is done for all DTs in the initial distribution, we evidently end up with an output distribution which is a standard normal distribution.

In the other two situations (middle and right panel in Figure S2A), the choice of the new DT_{zscore} is based on the previous DT_{zscore}'s. As discussed above, this will give rise to a correlation between the input and output distribution.

In the most extreme case, both DT_{zscore} distributions are perfectly correlated (right panel in Figure S2A). Each new DT_{zscore} will now be the same as the previous DT_{zscore}. Here, analogous to the transfer function in the left panel of Figure S2A, the transfer function can now be seen as a normal distribution with a mean equal to the previous DT_{zscore} and an infinitely small variance.

Between these two boundary conditions (left and right panel in Figure S2A), we find a range of scenarios where the correlation between the input and output distributions is not perfect, but still higher than zero ($0<R^2<1$; middle panel in Figure S2A). In this case, it is less trivial what the transfer function should be. It is however clear that it should be intermediate to the two extremes represented in the left and right panel of Figure S2A. The transfer function will thus be a normal distribution with a mean which is dependent on the initial DT_{zscore} (one parameter will determine how close this mean is to the initial DT_{zscore}), and a SD between 0 and 1. Note that the closer the mean of the transfer function is set to 0, the higher the SD of the transfer function will need to be. This is necessary to ensure that the output distribution has a variance of 1 (as is the case for a standard normal

distribution). Of course, the higher the SD of the transfer, the lower the R^2 of the correlation will be. By tuning these two parameters, we can thus model all degrees of correlation (R^2) between the two distributions.

7.3.2 Exact mathematical solution

Mathematically, the problem shown in Figure S2A, can be described as follows. First, we start with a standard normal input distribution.

$$f_{input}(x) \sim \mathcal{N}\left(\mu_{input}, \sigma^2_{input}\right)$$

$$with\ \mu_{input} = 0;\ \sigma^2_{input} = 1 \quad \textbf{(3)}$$

Next, the input distribution needs to be transformed by multiplying each value in this distribution with $A \in [0, 1]$. This is analogous to the effect of the blue arrows in Figure S2A. Consequently, the variance will scale down as described in equation 4, and a new normal distribution (input-prime) will be generated.

$$f_{input'}(x) \sim \mathcal{N}\left(\mu_{input'}, \sigma^2_{input'}\right)$$

$$with\ \mu_{input'} = \mu_{input} = 0;\ \sigma^2_{input'} = \left(A \times \sigma_{input}\right)^2\ with\ A \in [0, 1] \quad \textbf{(4)}$$

The transfer function is defined as a normal distribution with a mean of 0 and a given SD (equation 5).

$$f_{transfer}(x) \sim \mathcal{N}\left(\mu_{transfer}, \sigma^2_{transfer}\right)$$

$$with\ \mu_{transfer} = 0 \quad \textbf{(5)}$$

We can now generate an output distribution, by taking the convolution product of the input-prime distribution and the transfer function (equation 6). This is mathematically equivalent to the scheme shown in Figure S2A, where the transfer function is applied individually to each value in the input distribution, hereby generating a distribution of output values. Since the initial distribution and the transfer function are both normal distributions with a given mean and variance, the convolution of the two functions will again give rise to a normal distribution. The mean and variance of this resulting distribution are the sum of respectively the means and variances of the convolved functions [S6]. This is represented in equation 6 and 7.

$$f_{output}(x) = f_{input'}(x) * f_{transfer}(x) = \int_{-\infty}^{\infty} f_{input'}(\tau) \times f_{transfer}(x - \tau) \times d\tau \quad \textbf{(6)}$$

$$f_{output}(x) \sim \mathcal{N}\left(\mu_{output}, \sigma^2_{output}\right)$$

$$with\ \mu_{output} = \mu_{input'} + \mu_{transfer};\ \sigma^2_{output} = \sigma^2_{input'} + \sigma^2_{transfer} \quad \textbf{(7)}$$

As mentioned earlier, the output distribution needs to be a standard normal distribution, which means that $\mu_{output} = 0$ and $\sigma^2_{output} = 1$. Using equations 4-7, we can now calculate the equation relating our two unknown parameters A and $\sigma_{transfer}$.

$$\sigma^2_{transfer} = \sigma^2_{output} - \sigma^2_{input} \times A^2 = 1 - A^2 \quad \textbf{(8)}$$

Note the similarity to the different scenarios represented in Figure S2A. In the left panel, all values in the input distribution are first set to 0 ($A = 0$), and new values are then picked using a transfer function with $\sigma_{transfer} = 1$. In the right panel, the new output values are simply the same as the input distribution values ($A = 1; \sigma_{transfer} = 0$). And finally, in the middle panel, the intermediate scenario is shown. The variance of the input distribution is first scaled down by multiplying each value with $0 < A < 1$, and then new output values are chosen based on a $\sigma_{transfer}$ which depends on A (see equation 8).

7.3.3 Trial-and-error-based solution

In the previous section, we have described the mathematical relationships necessary to build DT correlations into our model. We will now test if these relationships still hold, when we move from idealized infinitely large distributions to distributions holding a finite number of DT_{zscore}'s (as in our simulations). Moreover, in this section we will deduce how different combinations of $\sigma_{transfer}$ and A lead to different degrees of correlation (R^2) between the input and output distributions.

In the following set of simulations, we started with a total of $N = 10\,000$ $DT_{input,zscore}$'s sampled from a standardized normal distribution.

$$DT_{input,zscore} = \sum_{i=1}^{N} DT_{i,input,zscore} \quad \textbf{(9)}$$

$$DT_{input,zscore} \sim \mathcal{N}(0,1) \quad \textbf{(10)}$$

Each of these values was then used as an input to generate corresponding $DT_{output,zscore}$ values using the following functions (which are mathematically equivalent to the equations 4-6 shown in section 7.3.2).

$$DT_{i,output,zscore} \sim \mathcal{N}\left(\mu_{transfer}, \sigma^2_{transfer}\right) \quad \textbf{(11)}$$

$$\mu_{transfer} = A \times DT_{i,input,zscore} \quad \textbf{(12)}$$

This results in a new set of $N = 10\,000$ $DT_{output,zscore}$ values.

$$DT_{output,zscore} = \sum_{i=1}^{N} DT_{i,output,zscore} \quad \textbf{(13)}$$

We then explored the parameter space for $\sigma_{transfer}$ and A by varying them independently from 0 to 1.1 in steps of 0.01. The outcome of these simulations was evaluated by considering the mean, SD and the normality of the resulting $DT_{output,zscore}$ distributions. Finally, we also kept track of the correlation between the $DT_{input,zscore}$ and $DT_{output,zscore}$ values (R^2). The results of these simulations are shown in Figure S2B.

The first prerequisite for finding the desired relationship between $\sigma_{transfer}$ and A, is that the resulting $DT_{output,zscore}$ distributions have a mean of 0 and a SD of 1 (as standard normal distributions). When we make a scatter plot of the parameters leading to a mean of 0 ($\pm$0.03) and a SD of 1 ($\pm$0.005), we can see a quarter circular line appearing (Figure S2B). Inspired by equation 8, we can now fit the following function relating $\sigma_{transfer}$ and A.

$$\sigma^2_{transfer} = 1 - A^2 \quad \textbf{(14)}$$

As expected from our exact mathematical solution, equation 14 describes the relationship between A and $\sigma_{transfer}$ well (R^2=0.9901). As can be seen from Figure S2B (panel iii), the resulting $DT_{output,zscore}$ distributions only lead to a normal distribution across a narrow range in the parameter space. Fortunately, this range coincides with the isoline representing a mean of 0 and a SD of 1 in the $DT_{output,zscore}$ distributions. Specifically, we find that the gross majority (>94%) of the distributions can be considered as normal ($p>0.05$ using Kolmogorov-Smirnov test for normality). Finally, the relationship between the R^2-values and the parameter A could be derived from Figure S2B (panel iv). The relationship was found to be linear on a log-log plot and could thus be approximated using the following function (R^2=0.9979).

$$\log(R^2) = 2.038 \times \log(A) \quad \textbf{(15)}$$

Our defined functions (equations 11-12,14), thus allow us to incorporate the whole spectrum of DT correlations ($0<R^2<1$) into our stochastic model.

7.4 Incorporating the mathematical framework into our stochastic model

Now that we have acquired mathematical functions which allow us to model the DT relationships that we have observed (Figure S1B-G), we can incorporate these into our growth model. In order to do this, it is important that we keep track of the DT that was last assigned to each cell (or the corresponding z-score of this DT). This DT (or rather, its z-score) can then be used to assign new DTs to the cells within each lineage, based on the equations described above.

Figure S2C shows which DT relationships are considered in our model. First, we have observed that consecutive mother DTs are on average slightly correlated (Figure S1D; R^2=0.0153). Second, we have found a much stronger DT correlation between the mother and her most recently born daughter (Figure 2d; R^2=0.2517). Using the equations 14 and 15, we calculated the parameters $\sigma_{transfer}$ and A which were used in equations 11 and 12 to derive the z-score of the next DT to be assigned. Note that our dataset does not allow us to investigate DT correlations between the DT of a cell when it is still a daughter and its DT when it has become a mother. Therefore, we conservatively assume that there is no correlation between both (Figure S2C).

7.5 Population growth rate predictions using the stochastic model which takes into account epigenetic DT inheritance

Now that we have constructed a stochastic model which takes into account the observed DT correlations, we can use it to predict population growth rates. This will allow us to investigate the effect of epigenetic DT inheritance on population growth rates. The model can be run using two sets of parameters. First, the average DT correlations can be used to calculate the parameters for the model. In this case, all experiments will have the same average level of epigenetic DT inheritance (Figure S5A). Second, the DT correlations can be calculated individually per experiment, leading to different parameters for the model for each experiment (Figure S5C). These two approaches are expected to lead to different outcomes, as the strength of the DT correlations can differ widely across different conditions (Figure 5A).

In the first approach, using the average epigenetic DT inheritance, we find a modest increase in population growth rates compared the purely stochastic model (Figure S5A). Furthermore, this increase is positively correlated to the variability present in the single-cell DT distributions (R^2 of growth rate increase vs. mother CoV= 0.8829; R^2 of growth rate increase vs. daughter CoV=0.6912; Figure S5B).

In the second approach, using the per-experiment individual epigenetic DT inheritance, we find on average that stronger increase in population growth rates compared to the purely stochastic model (Figure S5C). As expected, this increase is positively correlated to the strength of the DT correlations in each experiment (R^2 of growth rate increase vs. consecutive mother DT correlation = 0.3735; R^2 of growth rate increase vs. mother-daughter DT correlation=0.2570; Figure S5D).

8 Constructing a model which takes into account cell-cycle dependent single-cell growth

This study focuses on the effect of single-cell DT variability and epigenetic DT inheritance on the population growth rate. However, we also set out to investigate how these factors contribute to growth rate variability between different populations, more specifically at small population sizes when the variability in single-cell DTs might not yet be averaged out (Figure S6). For this purpose, we need to extend the model to take into account the volume increase during progression through the cell cycle. Indeed, at large population sizes, growth within single cells is unlikely to affect variability in population growth rates, since it is averaged out by the large number of cells present in the population. However, at small population sizes, growth within single cells is expected to play an important role in determining the population growth rate variability. Therefore, in the following section, we will extend the model to take into account single-cell growth as the cell progresses through the cell cycle.

Single-cell growth during progression through the cell cycle can be divided into two clearly distinct phases: (1) the Unbudded Period (UP), when the cell is not budded yet, and (2) the Budded Period (BP), from bud emergence until cell division. Using our collection of time-lapse growth records, we have measured these periods, by scoring the time of bud appearance in mother and daughter cells. For daughter cells, this was measured as shown in Figure 1A, and the results of these measurements are presented in Figure 2B. Additionally, we have scored the appearance of a bud in mother cells (not shown in Figure 1A) in a similar fashion, by scoring the bud appearance in the 1st and 2nd daughter cell after its first cell division (section 2.2 and Dataset S1). At this point, these cells have already undergone one cell division, and have thus effectively become a mother cell.

In Figure S6C-D, the mean and SD of the length of the UP are plotted against the total DT of both mother and daughter cells. Functions were fitted to these correlations and these were used in the growth simulations to determine the length of the mother Unbudded Period (mUP) and daughter Unbudded Period (dUP) for each cell. The relationships are shown in equation 16-21.

$$mUP = \mathcal{N}(mUP_{\text{mean}}, mUP_{\text{st.dev.}}) \quad \textbf{(16)}$$

$$mUP_{\text{mean}} = -0.1415 \times mDT^2 + 0.7879 \times mDT - 0.5397 \quad \textbf{(17)}$$

$$mUP_{\text{st.dev.}} = 0.6711 \times mDT - 0.6738 \quad \textbf{(18)}$$

$$dUP = \mathcal{N}(dUP_{\text{mean}}, dUP_{\text{st.dev.}}) \quad \textbf{(19)}$$

$$dUP_{\text{mean}} = 0.1686 \times dDT - 0.0268 \quad \textbf{(20)}$$

$$dUP_{\text{st.dev.}} = 0.2155 \times \ln(dDT) + 0.0545 \quad \textbf{(21)}$$

We need to simulate how much each bud has grown, and this will depend on the following three parameters: the length of the total DT, the length of the UP, and 'the time remaining for completion bud'. Therefore, during these simulations, we need to keep track of all these single-cell variables.

We decided to use a simple approximation of single-cell growth during the cell cycle. The assumption is that cells do not grow during their UP, and that growth proceeds in an exponential fashion during the BP. The mathematical function describing this is described by equation 22.

$$for\ "Time\ Remaining\ For\ Completion\ Bud" > DT - UP$$

$$"Fraction\ Of\ Bud\ Completed" = 0$$

$$for\ "Time\ Remaining\ For\ Completion\ Bud" < DT - UP$$

$$"Fraction\ Of\ Bud\ Completed" = 2 \times e^{\left(\frac{DT-UP-"Time\ Remaining\ For\ Completion\ Bud"}{DT-UP}\right)} - 1 \quad \textbf{(22)}$$

The size of the population is now expressed as the sum of all cells including the fractions of the bud which are already formed (equation 23). All microcolony growth rates in this study are total (not instantaneous) growth rates, as in equation 24.

$$pop.size = \sum_{i=1}^{N}(1 + "Fraction\ Of\ Bud\ Completed"_i) \quad \textbf{(23)}$$

$$microcolony\ growth\ rate\ (t_{final}) = \frac{ln(pop.size_{t.final}) - ln(pop.size_{t.initial})}{t_{final}} \quad \textbf{(24)}$$

As mentioned above, we are specifically interested in using this extension of the model to predict small population (microcolony) growth rate variability. To do this, we first grow (*in silico*) a large population of cells to steady-state (>100 000). Then we repeatedly sample single cells from this population which we allow to grow for a fixed amount of time. Finally, we calculate microcolony growth rates using equations 23 and 24 (for more details see section 4.4).

Since the purpose of this extended model is to investigate the effect of single-cell DT variability and epigenetic DT inheritance on the population growth rate variability, we decided to compare the predictions of our model with empirical measures of microcolony growth rate variability. We have measured population growth rates by tracking the area increase of N=19-62 microcolonies at three time points per experiment (Dataset S5; section 3.1). Therefore, from this data, we can extract two measures of microcolony growth rate variability per experiment. Next, this variability was simulated using the extended version of the model discussed here (with or without epigenetics). For each strain and condition, we drew the growth rates of 1000 random microcolonies from our simulation to predict variability (Dataset S5). In Figure S6A, the modeled microcolony growth rate variability was plotted against the empirical variability (both expressed as coefficient of variation). When running the model without epigenetic effects, the variability predictions correspond well to the empirical variability for most observations. However, when the empirical variability is high (>0.30), the model (with or without epigenetics) tends to strongly underestimate it.

Next, we used the Anderson-Darling (AD) test, a non-parametric test like the Kolmogorov-Smirnov Test, to check whether including epigenetic inheritance in the model improves the predictions of microcolony growth rate variability. Since microcolony growth rate are consistently underestimating actual growth rates (see above and Figure S3), all microcolony growth rates were normalized to within-distribution medians, to allow comparison of the variability. For most our experiments, the AD test cannot detect a significant difference when epigenetic inheritance in included in the model (Bonferroni corrected p-value>0.05; Dataset S5). Further analysis shows that for strain/condition combinations that display little epigenetic DT inheritance (mother-daughter DT correlation $R^2 < 0.35$), the model comparably matched the empirical data, independent of whether or not epigenetic effects were included (Figure S6B). However, for strains and conditions that demonstrate a high degree of epigenetic DT inheritance (mother-daughter DT correlation > 0.35), including epigenetics in the model made its predictions more closely matches the empirical variability (Figure S6B).

9 Validity of using the experimentally measured DT distributions to extract parameters for the model

Our model is most often used to predict population growth rates from experimentally measured single-cell parameters (mean and SD of mother and daughter DTs). However, these experimentally measured DT distributions do not necessarily correspond to the DT distributions that are sampled from to assign DTs to newly formed individuals. Indeed, Tănase-Nicola & ten Wolde (2008) have shown that the population average of a quantity (measured at a certain point in time) can be different from its time average (when tracking this quantity in a single-cell and its descendents), especially when fluctuations of this quantity influence growth rate [7]. Therefore, we investigated whether there is a discrepancy between the model's parameters and the experimentally measured values that are supposed to represent them.

To avoid confusion, we will first define the three different DT distributions that are relevant in this discussion.

1) **The (model-)sampling distribution:** The DT distribution from which the model samples when it assigns a new DT (to newly-born cells or cells that have just completed a previous doubling).

2) **The steady-state distribution:** The population DT distribution at any instant during steady-state growth. More specifically, this distribution can be measured (during simulations) by randomly sampling cells from one time point during exponential growth, and extracting the DTs that were last assigned to them.

3) **The experimental distribution:** The DT distribution that was measured experimentally in this study, by allowing the cells to finish the cell cycle that they were going through at the start of the movie, and following their subsequent growth as shown in Figure 1A. Similarly, these experimental distributions can be extracted during simulations by assigning new DTs to all cells present at a fixed time point.

In Figure S4D-E, we compare the three types of DT distributions described above for the model with and without epigenetics. The parameters used for these simulations were those as measured for BY/s288c growing in palatinose, a strain/condition that shows high DT noise and high epigenetic DT inheritance (Dataset S1).

The model-sampling distribution and the steady-state distribution are generally not the same under the stochastic model (Figure S4D), as the steady-state distribution is enriched for cells that were assigned a long DT (slowly-growing cells). This is because the fast-growing cells finish their doubling faster than the slow-growing cells, but the chance of being assigned a short or long DT are equal. Importantly, under the model assuming epigenetics, the fast-growing cells are more likely to give rise to fast-growing progeny, so the differences between the steady-state and model-sampling distributions become less pronounced (Figure S4E).

It is important to note that the experimental DT distributions are different from the steady-state distributions, because we do not measure the DTs at one instantaneous point in time. Instead, after spotting an aliquot of an exponentially growing culture on an agar gel, we randomly follow single-cells and only start scoring DTs after they finish their current cell cycle (Figure 1 and Movie S1). Under the purely stochastic model, newly-assigned DTs are independent of previously assigned DTs, and therefore the experimental distribution by definition equals the model-sampling distribution. When taking into account epigenetic inheritance, the experimental distribution has on average slightly more slow-growing cells, but not as much as the steady-state distribution (Figure S4D-E).

In conclusion, when predicting population growth rates from single-cell DT measurements, the experimentally measured DT distributions can be readily used to extract parameters for the purely stochastic model ($model_{stoch,no\text{-}epig}$). However, caution should be taken when predicting population growth rates from single-cell DT measurements using the model which assumes epigenetic effects ($model_{stoch,epig}$). As the epigenetic DT correlations increase, the experimental DT distribution will contain increasingly more slow-growing cells compared to the model-sampling DT distribution, which could lead to an underestimation of population growth rates. Importantly though, this does not affect the general result that increasing epigenetic DT inheritance (while keeping the DT noise and mean constant) leads to an increase in population growth rate (Figures 5B,C and S5).

10 Intuitive understanding of the effect of DT noise and epigenetic DT inheritance on the population growth rate

10.1 General

In these following sections, we try to provide the reader with an intuitive understanding of the effect that stochastic DT variability and epigenetic DT correlations have on the population growth rate.

First, we derive equations that allow the exact prediction of the (instantaneous) population growth rate when the size and growth rates of the subpopulations (or cells) are known (section 10.2). We show that the instantaneous growth rate is the weighted arithmetic mean of the growth rates of the subpopulations (cells), weighted by the size of the subpopulations (cells).

In section 10.3, we extend this equation to fit the parameters used in our stochastic model (the last DT assigned to each cell, and the time remaining until completion of the current cell cycle). We identify three factors that individually could lead to either an increase or decrease in population growth rate, when the DT noise in the model-sampling distribution (see section 9) is increased. These factors are 1) the growth rate associated with a certain DT, 2) the steady-state proportion of cells that have been assigned a certain DT and 3) the steady-state distribution of the progression through the cell cycle within these cells. We show that factor 1 and 3 theoretically could increase the population growth rate, while factor 2 theoretically could decrease it.

10.2 The instantaneous population growth rate in function of the sizes and growth rates of N subpopulations

Assume a population consisting of N different subpopulations. For simplicity, we will refer to these subpopulations as cells, and assume that their size=1 at time t=0. Each i-th cell grows at a fixed growth rate μ_i which can be different for each cell but stays constant through time. We will first derive what the average population growth rate is from time=0 until time=t.

The average growth rate of the whole population after time = t is given by the following formula.

$$\mu_{total}(t) \equiv \frac{ln\left(\frac{N_{total}(t)}{N_{total}(0)}\right)}{t} \quad \text{with } N_{total}(t) \text{ representing the total size of the population} \qquad \textbf{(25)}$$

At time=0 there are N different cells of size=1, so:

$$N_{total}(0) = \sum_{i=1}^{N} size_i(0) = N \qquad \textbf{(26)}$$

The total number of cells at time t, $N_{total}(t)$, is then given by:

$$N_{total}(t) = \sum_{i=1}^{N} size_i(0) * exp(\mu_i * t) = \sum_{i=1}^{N} exp(\mu_i * t) \qquad \textbf{(27)}$$

The combination of equations 25-27 gives:

$$\mu_{total}(t) = \frac{\ln\left(\sum_{i=1}^{N} exp(\mu_i * t)\right) - ln(N)}{t} \qquad \textbf{(28)}$$

To know what the instantaneous growth rate is of the population at time = 0, we have to calculate the limit of this function at time = 0.

$$\lim_{t\to 0} \mu_{total}(t) = \frac{\ln\left(\sum_{i=1}^{N} exp(\mu_i * 0)\right) - ln(N)}{0} = \frac{\ln(N) - ln(N)}{0} = \frac{0}{0} \qquad \textbf{(29)}$$

This limit's solution is now in an indeterminate form, but can be calculated through L'Hôpital's rule if the some constraints are fulfilled.

L'Hôpital's rule states that if

$$\lim_{t\to 0} f(t) = \lim_{t\to 0} g(t) = 0 \text{ or } \pm\infty \text{ and } \lim_{t\to 0} \frac{f\prime(t)}{g\prime(t)} \text{ exists} \qquad \textbf{(30)}$$

then

$$\lim_{t\to 0} \frac{f(t)}{g(t)} = \lim_{t\to 0} \frac{f\prime(t)}{g\prime(t)}. \quad \textbf{(31)}$$

Applying L'Hôpital's rule using $f(t) = \ln(\sum_{i=1}^{N} exp(\mu_i * t)) - ln(N)$ and $g(t) = t$ gives:

$$\lim_{t\to 0} \mu_{total}(t) = \frac{D\{\ln(\sum_{i=1}^{N} exp(\mu_i * t)) - ln(N)\}}{D\{t\}} \quad \textbf{(32)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{D\{\ln(\sum_{i=1}^{N} exp(\mu_i * t))\} - D\{ln(N)\}}{D\{t\}} \quad \textbf{(33)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{D\{\ln(\sum_{i=1}^{N} exp(\mu_i * t))\} - 0}{1} \quad \textbf{(34)}$$

$$\lim_{t\to 0} \mu_{total}(t) = D\{\ln(\sum_{i=1}^{N} exp(\mu_i * t))\} \quad \textbf{(35)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{1}{\sum_{i=1}^{N} exp(\mu_i * t)} D\{(\sum_{i=1}^{N} exp(\mu_i * t))\} \quad \textbf{(36)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{1}{\sum_{i=1}^{N} exp(\mu_i * t)} * (\sum_{i=1}^{N} D\{exp(\mu_i * t)\}) \quad \textbf{(37)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{1}{\sum_{i=1}^{N} exp(\mu_i * t)} * (\sum_{i=1}^{N} exp(\mu_i * t) * D\{\mu_i * t\}) \quad \textbf{(38)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{1}{\sum_{i=1}^{N} exp(\mu_i * t)} * (\sum_{i=1}^{N} exp(\mu_i * t) * \mu_i\}) \quad \textbf{(39)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{1}{\sum_{i=1}^{N} exp(\mu_i * 0)} * (\sum_{i=1}^{N} exp(\mu_i * 0) * \mu_i\}) \quad \textbf{(40)}$$

$$\lim_{t\to 0} \mu_{total}(t) = \frac{1}{N} * (\sum_{i=1}^{N} 1 * \mu_i\}) \quad \textbf{(41)}$$

$$\lim_{t\to 0} \mu_{total}(t) = arithmetic\ mean\ of\ \mu_i\ (with\ i = 1,2,3,..,N) \quad \textbf{(42)}$$

This result is based on the assumption that the initial size of each cell equals 1. The result can easily be extended to the situation where each i-th cell has an initial size of $size_i(t = 0)$ and the total initial size of the population equals $\sum_{i=1}^{N} size_i(0)$.

$$\lim_{t\to 0} \mu_{total}(t) = \frac{1}{\sum_{i=1}^{N} size_i(0)} * (\sum_{i=1}^{N} size_i(0) * \mu_i\})$$

$$= weighted\ arithmetic\ mean\ of\ \mu_i, weighted\ by\ size_i(with\ i = 1,2,3,..,N) \quad \textbf{(43)}$$

In conclusion, the instantaneous population growth rate of a population is the weighted arithmetic mean of the growth rates of its cells (more generally, subpopulations), weighted by the sizes of its cells (subpopulations).

10.3 The steady-state growth rate in function of the variables used in our stochastic model

In the previous section, we derived an equation that allows the calculation of the instantaneous population growth rate, based on the growth rates and sizes of all single-cells within that population. To make these equations compatible with our model, we first need to define the size and growth rate of each cell, based on the two single-cell parameters that can fully describe the state of each cell in our stochastic model. The first parameter is the DT assigned to each cell DT_i, and the second parameter is the time remaining for completion of its doubling τ_i.

In our model, each i-th cell start out at a size of 1, and grows exponentially at a rate of $\ln(2)/DT_i$ until it reaches a size of 2. The size and the growth rate of each cell can thus be represented using the following formula.

$$size_i = 2^{\left(\frac{DT_i - "Time\ Remaining\ For\ Completion\ Bud"}{DT_i}\right)} = 2^{\left(\frac{DT_i - \tau_i}{DT_i}\right)} \quad \textbf{(44)}$$

$$\mu_i = \frac{\ln(2)}{DT_i} \quad \textbf{(45)}$$

Combing equations 43 until 45 gives the instantaneous growth rate in terms of the parameters used in our model.

$$\mu_{instantaneous}(t) = \frac{1}{\sum_{i=1}^{N} 2^{\left(\frac{DT_i - \tau_i}{DT_i}\right)}} * \left(\sum_{i=1}^{N} 2^{\left(\frac{DT_i - \tau_i}{DT_i}\right)} * \frac{\ln(2)}{DT_i} \right) \quad \textbf{(46)}$$

Since the parameters used in our model represent continuous variables, we will now derive the same equation, but using integrals. For this, we will define the function $p(DT, \tau, t)$, which describes the probability to find a cell in a state given a combination of DT and τ, at time = t.

$$p(DT, \tau, t) = \ probability\ to\ find\ a\ cell\ in\ a\ state\ given\ DT\ and\ \tau$$

$$with \quad \int_0^\infty \int_0^\infty p(DT, \tau, t) * dDT * d\tau = 1 \qquad for\ all\ values\ of\ t \quad \textbf{(47)}$$

Given equation 47, the instantaneous growth rate in terms of the parameters used in our model can be written in the integral form.

$$\mu_{instantaneous}(t) = \frac{1}{\int_0^\infty \int_0^\infty p(DT,\tau,t) * 2^{\left(\frac{DT-\tau}{DT}\right)} * dDT * d\tau} * \int_0^\infty \int_0^\infty p(DT, \tau, t) * 2^{\left(\frac{DT-\tau}{DT}\right)} * \frac{ln(2)}{DT} * dDT * d\tau \quad \textbf{(48)}$$

During steady-state growth, the probability $p(DT, \tau, t)$ function and the population growth rate will have reached a stable distribution. Therefore, the steady-state probability distribution ($\equiv p_{ss}(DT, \tau)$) and the steady-state population growth rate ($\equiv \mu_{ss}$) are not dependent on time.

$$\mu_{ss} = \frac{1}{\int_0^\infty \int_0^\infty p_{ss}(DT,\tau) * 2^{\left(\frac{DT-\tau}{DT}\right)} * dDT * d\tau} * \int_0^\infty \int_0^\infty p_{ss}(DT, \tau) * 2^{\left(\frac{DT-\tau}{DT}\right)} * \frac{ln(2)}{DT} * dDT * d\tau \quad \textbf{(49)}$$

The denominator in equation 50 can be thought of as the arithmetic mean of the cell sizes, and will be summarized as 'MeanCellSize'.

$$\mu_{ss} = \frac{1}{MeanCellSize} \int_0^\infty \int_0^\infty p_{ss}(DT, \tau) * 2^{\left(\frac{DT-\tau}{DT}\right)} * \frac{ln(2)}{DT} * dDT * d\tau \quad \textbf{(50)}$$

10.4 Intuitive understanding of the population growth rate effect of DT noise in the model$_{\text{stoch,no-epig}}$

In order to intuitively understand how DT noise can affect the steady-state population growth, it can be useful to divide the population into groups of cells that were assigned the same DT, and consider how much each group contributes to the total population growth.

$$\mu_{ss} \equiv \int_0^\infty GrowthRateContribution\ (DT) * dDT \quad \textbf{(51)}$$

Given equations 50 and 51, the growth rate contribution at a certain DT will be given by the following equation.

$$GrowthRateContribution\ \ (DT) = \frac{1}{MeanCellSize} * \frac{ln(2)}{DT} * \int_0^\infty p_{ss}(DT, \tau) * 2^{\left(\frac{DT-\tau}{DT}\right)} * d\tau \quad \textbf{(52)}$$

The ratio $\frac{ln(2)}{DT}$ within the equation represents the growth rate of the cells, while the right part $\int_0^\infty p_{ss}(DT, \tau) * 2^{\left(\frac{DT-\tau}{DT}\right)} * d\tau$ represents the proportion of the cells that have been assigned this DT but weighted by the size of these cells.

In order to understand how population growth rate is affected by adding (Gaussian) noise to the sampling DT distribution, while keeping the mean DT μ_{DT} the same, one can compare the growth rate contribution at the DT = $\mu_{DT} - C$ with the growth rate contribution at the DT = $\mu_{DT} + C$ (with C being a positive constant). If the average growth rate contribution of DT = $\mu_{DT} - C$ and DT = $\mu_{DT} + C$ is higher than the growth rate contribution at DT = μ_{DT}, the population growth rate will increase, and vice versa.

The effect of changing the DT on the single-cell growth rate is straightforward since it is given by equation $\frac{ln(2)}{DT}$. The longer the DT that is assigned to a cell, the slower it will grow. However, since this is a convex function, the

growth rate increase of cells that were assigned a DT = $\mu_{DT} - C$, is higher than the growth rate decrease of cells that were assigned a DT = $\mu_{DT} + C$. In other words:

$$\frac{\ln(2)}{DT-c} - \frac{\ln(2)}{DT} > \frac{\ln(2)}{DT} - \frac{\ln(2)}{DT+c} \quad \textbf{(53)}$$

Therefore, without considering the proportions and sizes of cells during steady-state in equation 52, we can expect that increasing the (Gaussian) noise in the model-sampling DT distribution would increase the population growth rate, since the growth rate contribution at DT = $\mu_{DT} - C$ would increase more, than the growth rate contribution at DT = $\mu_{DT} + C$ would decrease. This mathematical consequence is more generally known as the Jensen's inequality, which states that the mean of a set of values that have undergone a convex transformation f(x), is equal to or higher than the same transformation of the mean of these values (or mean[f(x)] ≥ f(mean[x])).

How adding DT noise affects the steady-state distribution is more complicated. But based on equation 52, we can identify two factors herein that will have an effect on the population growth rate: the steady-state proportion of cells that were assigned a DT, and the size distribution within these groups of cells given a certain DT.

It can intuitively be understood that increasing the DT noise in the sampling distribution (while keeping the DT mean constant) will lead to a higher proportion of cells with long DTs during steady-state. Indeed, while each cell has a similar chance of being assigned a long or a short DT, the cells with a short DT will finish their cell cycle more quickly. The net effect is that the steady-state population will be enriched with long DTs compared to the model-sampling distribution (Figure S4D). The higher the DT noise in the sampling distribution, the bigger this effect will be. This will reduce the population growth rate contribution of the cells which were assigned short DTs.

Finally, increasing the DT noise in the sampling distribution (while keeping the mean DT constant) will increase on average the size of cells with short DT. This is because, equal proportions of cells with long and short DTs are assigned at each instant, but cells with a long DT progress increase in size more slowly. This will increase the population growth rate contribution of the cells that were assigned short DTs.

In summary, we can identify three factors that influence the population growth rate under stochastic DT variability. In comparison to cells with long DTs, cells that have been assigned short DTs will: 1) grow faster, 2) become less abundant during steady-state, and 3) have larger sizes during steady-state. Factors 1 and 3 will increase the growth contribution of cells with short DTs, while factor 2 will decrease it. The net effect of these three factors is difficult to predict a priori, but our $\text{model}_{\text{stoch,no-epig}}$ shows that the combined effect is an increase in population growth rate (Figure 4 and S4).

10.5 Population growth rate effect of DT noise in the $\text{model}_{\text{stoch,epig}}$

When the model is run assuming epigenetic DT correlations ($\text{model}_{\text{stoch,epig}}$), the population growth rate will be dependent on the same three factors as described in section 10.4. In this framework, the predicted population growth rate increase due to epigenetic DT correlations, can be explained by its effect on the steady-state distribution $p_{ss}(DT, \sigma)$. Similar as in the $\text{model}_{\text{stoch,no-epig}}$, cells that have been assigned short DTs will become depleted from the steady-state population. However, the effect is less pronounced because these cells now have a higher chance to be assigned a short DT again. This will reduce the differences in abundance and average size that exist during steady-state between cells that have been assigned long and short DTs. While reducing the differences in abundance could lead to an increase in population growth rate, reducing the differences in average size could lead to a decrease. The net effect is difficult to predict a priori, but our $\text{model}_{\text{stoch,epig}}$ predictions show that the combined effect is an increase in population growth rate (Figure 5 and S5).

Supplemental Experimental Procedures

S1. Brachmann, C., Davies, A., Cost, G., Caputo, E., Li, J., Hieter, P., and Boeke, J. (1998). Deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast *14*, 115–132.

S2. Brown, C., Murray, A. W., and Verstrepen, K. J. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. Curr. Biol. *20*, 895–903.

S3. Dimitrov, L. N., Brem, R. B., Kruglyak, L., and Gottschling, D. E. (2009). Polymorphisms in multiple genes contribute to the spontaneous mitochondrial genome instability of *Saccharomyces cerevisiae* S288C strains. Genetics *183*, 365–383.

S4. New, A. M., Cerulus, B., Govers, S. K., Perez-Samper, G., Zhu, B., Boogmans, S., Xavier, J. B., and Verstrepen, K. J. (2014). Different levels of catabolite repression optimize growth in stable and variable environments. PLoS Biol. *1*, e1001764.

S5. Rosin, D., Hornung, G., Tirosh, I., Gispan, A., and Barkai, N. (2012). Promoter nucleosome organization shapes the evolution of gene expression. PLoS Genet. *8*, e1002579.

S6. Bromiley, P. A. (2003). Products and Convolutions of Gaussian Probability Density Functions Density Functions. Tina-Vision Memo *3*.

S7. Tănase-Nicola, S., and ten Wolde, P. R. (2008). Regulatory control and the costs and benefits of biochemical noise. PLoS Comput. Biol. *4*, e1000125.

Supplemental Dataset Legends

Dataset S1. Related to Figure 1,2 and 6. All single-cell DT measurements obtained using time-lapse microscopy. This includes firstly all measurements from the dataset used throughout this study (n=41 experiment; see Figures 1,2), and secondly the measurements used exclusively for Figure 6. The latter measurements are flagged under the column 'Used exclusively in Fig 6'.

(A) All measured single-cell DTs and (un)budded periods. All DTs and daughter (un)budded periods were scored as shown in Figure 1A. Mother (un)budded periods were scored for 21 out of 41 experiments (not shown in Figure 1A). Missing observations are shown by "NA", and represent individual observations that could not be scored (due to overcrowding) or observations that were not scored across a whole experiment (see Supplemental Experimental Procedures).

(B) Mean, SD, CoV, skewness and kurtosis of mother and daughter DT distributions. Per experiment, the mean, SD, CoV, skewness and kurtosis were calculated for the mother and daughter DT distributions (using the data from (A)). These calculations were done using a bootstrapping (resampling) approach (see Supplemental Experimental Procedures). For each parameter, this led to a 'bootstrapped mean' and a ' bootstrapped SD'. These 'bootstrapped SDs' were used to generate error bars in Figure 2A,C, Figure 4A and Figure S4A.

(C) Epigenetic DT inheritance per experiment. R represents the Pearson's correlation coefficient for the correlations between DTs of the cells within microcolonies. The two DT correlations which were considered in the model were calculated for each experiment individually: 1) the correlation between consecutive mother DTs (mDT_1 vs. mDT_2) and 2) the correlation between DTs of a mother and its most recently born daughter ($mDT_{1/2}$ vs. $dDT_{1/2}$). The distribution of R^2 for both correlations across our dataset is shown in Figure 5A. N represents the number of data points used to calculate R and R^2. All these parameters were calculated using the DT data from (A).

(D) Cumulative mother and daughter DT distributions per experiment. These cumulative DT distributions are represented in Figure 1B.

Dataset S2. Related to Figure 3. Testing Goodness-of-Fit of Normal, Gamma, Weibull, Log-logistic, Log-normal and Exponential probability distributions for observed DT data.

(A) Fitting probability distributions to DT data. All mother and daughter DT measurements (Dataset S1) were grouped per experiment, and Normal, Gamma, Weibull, Log-logistic, Log-normal and Exponential distributions were fitted to them using a bootstrapping (resampling) approach (see Supplemental Experimental Procedures). In short, the DT data was resampled 1000 times with replacement at full population size and distributions were fitted using the fitdist() function in MATLAB. For each of the fitted distribution's parameters, this led to a 'bootstrapped mean' and 'bootstrapped SD'.

(B) Goodness-of-fit using Anderson-Darling Test. The goodness-of-fit of the distributions fitted to the bootstrapped DT data in (A) is tested using the Anderson-Darling test. For each bootstrapped dataset from (A), the goodness-of-fit was examined for the different fitted distributions (Normal, Gamma, Weibull, Log-logistic, Log-normal and Exponential), and p-values were calculated using the function adtest() in MATLAB. As the DT data was resampled 1000 times per experiment, this led to 1000 p-values per experiment per distribution. To assess whether a certain distribution gave a good fit with our experimental data, an average p-value was calculated as the mean of these 1000 'bootstrapped p-values'. Also see Supplemental Experimental Procedures.

Dataset S3. Related to Figure 3-5. Modeled and empirical population growth rates.

(A) Predicted population growth rates for the whole dataset using the deterministic and stochastic model (with and without epigenetic effects). This table contains all predicted population growth rates which were obtained using the models discussed in this study: 1) the deterministic model (Hartwell & Unger, 1977), 2) our stochastic model without epigenetics and 3) our stochastic model with epigenetics. All growth rate predictions of the model were obtained using a bootstrapping (resampling) approach

(Supplemental Experimental Procedures). For each of the predictions, this led to a 'bootstrapped mean' and 'bootstrapped SD'. Furthermore, the specific values of the single-cell parameters used in each model are listed. When running our model assuming epigenetics, two different sets of parameters were used: 1) 'Global epigenetic DT inheritance', using the average DT correlations measured on all DT data combined across all experiments and 2) 'Individual epigenetic DT inheritance', using the DT correlations measured for each experiment individually.

(B) Experimentally measured population growth rates on solid media. Microcolony growth rates were measured by tracking either the cell count (Figure 3B) or the area (Figure S3B) increase over time. Per time-lapse movie, a number of microcolonies were analyzed by measuring the area (or cell count) at three evenly spaced time-points throughout the movie. Next, the area (or cell count) of all the microcolonies at each timepoint was summed and the microcolony growth rate was calculated by fitting a linear model to the ln-transformed summed area (or cell count) plotted against time (Supplemental Experimental Procedures). Listed are the slope of this linear model (the microcolony growth rate), the R^2 of the correlation, the fold increase in area (or cell count) from the first time point to the last, and the number of microcolonies scored.

(C) Experimentally measured population growth rates in liquid media. These measurements were used in Figure 3C. Growth rates were measured in liquid culture using a standard colony counting assay (Supplemental Experimental Procedures).

(D) Predicted population growth rates using the empirical DT distributions, Gamma DT distributions and Pearson system DT distributions. In contrast to (A), the population growth rates under the $model_{stoch,no\text{-}epig}$ were predicted using other DT distributions than the Normal distribution. This data is shown in Figure S4A.

(E) Predicted population growth rates using different DT averages in the deterministic model. In (A), the deterministic model uses the arithmetic mean DT to estimate the population growth rate. Here, we show the population growth rate predictions when using different averages (the median and the arithmetic, geometric, harmonic, quadratic and cubic mean). These averages were taken directly on the DT measurements, or after converting them to 'growth rate'-values [using GR = ln(2)/DT] followed by conversion of the average back to a DT-value. This data was used to calculate the mean squared error and mean error compared to the stochastic model in Figure S4C.

Dataset S4. Related to Figure 1-6. Summary statistics of different measurements and simulations discussed in the manuscript.

This dataset provides a table with summary statistics of different measurements and simulations discussed in the manuscript. All of these data are also shown in Datasets S1 and S3, but are combined here in a minimal and easy-to-handle table that was used for statistical analysis. All DT-related summary statistics were obtained by bootstrapping, and re-calculation of a given statistic can be found in Dataset S1 (Supplemental Experimental Procedures). Model-related statistics were generated as described by repeated Monte Carlo simulations (Supplemental Experimental Procedures). Population-level growth rates were determined as described in the Supplemental Experimental Procedures.

Dataset S5. Related to Figure 3. Modeled and empirical microcolony growth rate variability

(A) Experimental and predicted population growth rates for the whole dataset. The microcolony growth rates measurements represented in this table were also used to calculate population growth rates on solid medium by tracking area increase (Supplemental Experimental Procedures). Per time-lapse movie, a number of microcolonies (19-62) were analyzed by measuring the area at three evenly spaced time-points throughout the movie. Microcolony growth rates were then calculated 1) from the initial to the intermediate time point and 2) from the initial to the final time point. To generate predictions of microcolony growth rates, for each strain and condition, we randomly selected the modeled growth rates of 1000 microcolonies from our simulation. This was done using the strictly stochastic and the stochastic-epigenetic model, which were both extended to take into account the cell-cycle dependent volume increase within single cells (Supplemental Experimental Procedures).

(B) Statistical analysis of accuracy of microcolony variability predictions using Anderson-Darling (AD) Test. Initially, all microcolony growth rates (from (A)) were normalized to within-distribution medians to be able to specifically compare the variability of the modeled and empirical microcolony

growth rates. Next, the predicted distributions were tested against the empirical distributions using adk.test() function in R. For each comparison, the AD-Test Statistic, the p-value and the Bonferroni-corrected p-value are listed. The parameter "Estimated Population Size" was calculated using the following formula: EstimatedPopulationSize = 1 * exp (microcolonyGrowthRate * timeElapsed) with microcolonyGrowthRate equaling the average microcolony growth rate.

Supplemental Movie Legends

Movie S1. Related to Figure 1. Measurement of single-cell doubling times of mother and daughter cells using time-lapse differential interference contrast (DIC) microscopy.

A time-lapse movie of BY/S288c growing in YP + 3% glucose is shown to indicate which events were scored during analysis. A subset of time-lapse pictures from this movie was used in Figure 1A (after cropping to show only one microcolony). The color-code that was used to indicate the mother (red), 1st daughter (blue), 2nd daughter (green) and 3th daughter (yellow) corresponds to the color-code used in Figure 1A. The different events that were scored during the analysis are displayed on-screen during the movie.